(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)特許出願公表番号
特表2005-505037
(P2005-505037A)

(43)公表日 平成17年2月17日(2005.2.17)

(51) Int.Cl.[7]	F I		テーマコード（参考）
G06F 13/10	G06F 13/10	340A	5B014
G06F 3/06	G06F 3/06	301A	5B065
G06F 12/00	G06F 3/06	301N	5B082
H04L 29/06	G06F 12/00	545A	5K034
	H04L 13/00	305B	

審査請求 未請求 予備審査請求 有 （全 171 頁）

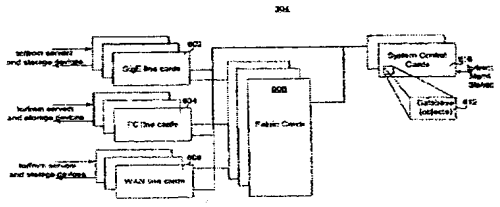
(21) 出願番号 特願2003-531344 (P2003-531344)
(86) (22) 出願日 平成14年9月27日 (2002.9.27)
(85) 翻訳文提出日 平成16年3月29日 (2004.3.29)
(86) 国際出願番号 PCT/US2002/030912
(87) 国際公開番号 WO2003/027877
(87) 国際公開日 平成15年4月3日 (2003.4.3)
(31) 優先権主張番号 60/325,704
(32) 優先日 平成13年9月28日 (2001.9.28)
(33) 優先権主張国 米国 (US)
(31) 優先権主張番号 10/051,415
(32) 優先日 平成14年1月18日 (2002.1.18)
(33) 優先権主張国 米国 (US)

(71) 出願人 504121830
マランティ ネットワークス インコーポレイテッド
アメリカ合衆国 カリフォルニア州 95134-2127 サン ホセ ザンカー ロード 3061 ビー
(74) 代理人 100082005
弁理士 熊倉 禎男
(74) 代理人 100067013
弁理士 大塚 文昭
(74) 代理人 100074228
弁理士 今城 俊夫
(74) 代理人 100086771
弁理士 西島 孝喜

最終頁に続く

(54) 【発明の名称】記憶システムにおけるプロトコル変換

(57) 【要約】
本発明の実施形態による記憶装置スイッチは、ＳＡＮの作成を可能にし、分散が容易で、中央管理が可能な拡張性の高いスイッチである。更に、この記憶装置スイッチは、世界的規模のインフラストラクチャの分散を可能にするので、記憶装置等のＳＡＮを実質的に世界のどこにでも配置することができる。更に、この記憶装置スイッチは、ｉＳＣＳＩ又はファイバ・チャンネルの両方を含むもの等の多重プロトコルＳＡＮを可能にし、データパケットを「ワイヤ速度」で処理する。更なるワイヤ速度処理を可能にするために、本発明によるスイッチは、ラインカードの各々に分散された「知能」を有し、ラインカードによって、パケットをデータ及び制御パケットに分類し、仮想化機能を実行し、プロトコル変換機能を実行する。更に、本発明によるスイッチは、ミラーリング、スナップショット、複製等のサーバレス記憶サービスを実行する。



【特許請求の範囲】

【請求項１】

データを記憶してアクセスするためのシステムに使用する方法であって、

（ａ）第１のプロトコルに従ってフォーマットされたパケットを、前記第１のプロトコルに従って作動する第１の装置から受信するステップと、

（ｂ）前記パケットを第２のプロトコルに従ってフォーマットされたパケットに変換するステップと、

（ｃ）前記パケットを前記第２のプロトコルに従って作動する第２の装置に送信するステップと、

を含み、前記ステップ（ａ）から（ｃ）は、前記パケットのバッファリングを行うことなく実行されることを特徴とする方法。

【請求項２】

前記ステップ（ａ）から（ｃ）は、ワイヤ速度で実行されることを特徴とする請求項１に記載の方法。

【請求項３】

前記変換ステップは、前記第１のプロトコルに従ってフォーマットされた前記パケットのフィールドの少なくとも一部を、前記第２のプロトコルに従ってフォーマットされた前記パケットにマッピングするステップを含むことを特徴とする請求項１に記載の方法。

【請求項４】

前記変換ステップは、前記第１のプロトコルに従ってフォーマットされた前記パケットからマッピングされていない、前記第２のプロトコルに従ってフォーマットされた前記パケットの任意のフィールドに情報を追加するステップを更に含むことを特徴とする請求項３に記載の方法。

【請求項５】

前記第１のプロトコルはｉＳＣＳＩであり、前記第２のプロトコルはファイバ・チャンネルであることを特徴とする請求項１に記載の方法。

【請求項６】

前記第１のプロトコルは、ＴＣＰに組み込まれたｉＳＣＳＩであることを特徴と

する請求項５に記載の方法。

【請求項７】

前記第１のプロトコルはファイバ・チャンネルであり、前記第２のプロトコルはｉＳＣＳＩであることを特徴とする請求項１に記載の方法。

【請求項８】

前記第２のプロトコルは、ＴＣＰに組み込まれたｉＳＣＳＩであることを特徴とする請求項７に記載の方法。

【請求項９】

前記第１の装置はイニシエータであり、前記第２の装置はターゲットであることを特徴とする請求項１に記載の方法。

【請求項１０】

前記第１の装置はターゲットであり、前記第２の装置はイニシエータであることを特徴とする請求項１に記載の方法。

【請求項１１】

データを記憶してアクセスするためのシステムに使用する方法であって、

（ａ）第１のプロトコルに従ってフォーマットされたパケットを、前記第１のプロトコルに従って作動する第１の装置から受信するステップと、

（ｂ）前記パケットを第２のプロトコルに従ってフォーマットされたパケットに変換するステップと、

（ｃ）前記パケットを前記第２のプロトコルに従って作動する第２の装置に送信するステップと、

を含み、前記ステップ（ａ）から（ｃ）は、ワイヤ速度で実行されることを特徴とする方法。

【請求項１２】

データを記憶してアクセスするためのシステムに使用する方法であって、

第１の認識プロトコルに従ってフォーマットされたパケットを受信するステップと、

前記パケットのバッファリングを行うことなく、前記パケットのフィールドを第２のプロトコルに従ってフォーマットされた新しいパケットにマッピングするス

テップと、

を含むことを特徴とする方法。

【請求項１３】

前記受信ステップ及び前記マッピングステップは、ワイヤ速度で実行されることを特徴とする請求項１２に記載の方法。

【請求項１４】

前記マッピングステップは、バッファリングを行うことなく、且つＣＰＵリソースを使用することなく、前記パケットの前記フィールドを、第２のプロトコルに従ってフォーマットされた新しいパケットにマッピングするステップを含むことを特徴とする請求項１２に記載の方法。

【請求項１５】

前記第１のプロトコルはｉＳＣＳＩであり、前記第２のプロトコルはファイバ・チャンネルであることを特徴とする請求項１２に記載の方法。

【請求項１６】

前記第１のプロトコルは、ＴＣＰに組み込まれたｉＳＣＳＩであることを特徴とする請求項１５に記載の方法。

【請求項１７】

前記第１のプロトコルはファイバ・チャンネルであり、前記第２のプロトコルはｉＳＣＳＩであることを特徴とする請求項１２に記載の方法。

【請求項１８】

前記第２のプロトコルは、ＴＣＰに組み込まれたｉＳＣＳＩであることを特徴とする請求項１７に記載の方法。

【請求項１９】

前記パケットは、ＳＣＳＩコマンド記述子ブロック（ＣＤＢ）を含むことを特徴とする請求項１２に記載の方法。

【請求項２０】

データを記憶してアクセスするためのシステムに使用する方法であって、

ｉＳＣＳＩ又はファイバ・チャンネルプロトコルの１つに従ってフォーマットされたパケットを第１の装置から受信するステップと、

前記パケットが前記ｉＳＣＳＩプロトコルに従ってフォーマットされている場合、前記ｉＳＣＳＩプロトコルに従ってフォーマットされた前記パケットからのフィールドの少なくとも一部をファイバ・チャンネルパケットにマッピングすることによって、前記パケットを前記ファイバ・チャンネルパケットに変換するステップと、

前記パケットが前記ファイバ・チャンネルプロトコルに従ってフォーマットされている場合、前記ファイバ・チャンネルプロトコルに従ってフォーマットされた前記パケットからのフィールドの一部を前記ｉＳＣＳＩパケットにマッピングすることによって、前記パケットを前記ｉＳＣＳＩパケットに変換するステップと、

前記パケットを変換された状態で第２の装置に送信するステップと、

を含み、前記ステップの全ては、バッファリングを行うことなく実行されることを特徴とする方法。

【請求項２１】

前記ステップの全ては、ワイヤ速度で実行されることを特徴とする請求項２０に記載の方法。

【請求項２２】

前記第１の装置はイニシエータであり、前記第２の装置はターゲットであることを特徴とする請求項２０に記載の方法。

【請求項２３】

前記第１の装置はターゲットであり、前記第２の装置はイニシエータであることを特徴とする請求項２０に記載の方法。

【請求項２４】

前記受信されたパケットはｉＳＣＳＩコマンドＰＤＵであり、前記変換された状態のパケットはファイバ・チャンネルコマンドＩＵであることを特徴とする請求項２０に記載の方法。

【請求項２５】

前記受信されたパケットはファイバ・チャンネルＸＦＲ＿ＲＤＹ　ＩＵであり、前記変換された状態のパケットはｉＳＣＳＩ　Ｒ２Ｔ　ＰＤＵであることを特徴

とする請求項２０に記載の方法。

【請求項２６】

前記受信されたパケットはｉＳＣＳＩ書込みデータＰＤＵであり、前記変換された状態のパケットはファイバ・チャンネルデータＩＵであることを特徴とする請求項２０に記載の方法。

【請求項２７】

前記受信されたパケットはファイバ・チャンネルデータＩＵであり、前記変換された状態のパケットはｉＳＣＳＩ読出しデータＰＤＵであることを特徴とする請求項２０に記載の方法。

【請求項２８】

前記受信されたパケットはファイバ・チャンネル応答ＩＵであり、前記変換された状態のパケットはｉＳＣＳＩ応答ＰＤＵであることを特徴とする請求項２０に記載の方法。

【請求項２９】

前記受信されたパケットはファイバ・チャンネルコマンドＩＵであり、前記変換された状態のパケットはｉＳＣＳＩコマンドＰＤＵであることを特徴とする請求項２０に記載の方法。

【請求項３０】

前記受信されたパケットはｉＳＣＳＩ　Ｒ２Ｔ　ＰＤＵであり、前記変換された状態のパケットはファイバ・チャンネルＸＦＲ＿ＲＤＹ　ＩＵであることを特徴とする請求項２０に記載の方法。

【請求項３１】

前記受信されたパケットはファイバ・チャンネルデータＩＵであり、前記変換された状態のパケットはｉＳＣＳＩ書込みデータＰＤＵであることを特徴とする請求項２０に記載の方法。

【請求項３２】

前記受信されたパケットはｉＳＣＳＩ読出しデータＰＤＵであり、前記変換された状態のパケットはファイバ・チャンネルデータＩＵであることを特徴とする請求項２０に記載の方法。

【請求項３３】

前記受信されたパケットはｉＳＣＳＩ応答ＰＤＵであり、前記変換された状態のパケットはファイバ・チャンネル応答ＩＵであることを特徴とする請求項２０に記載の方法。

【請求項３４】

データを記憶してアクセスするためのシステムに使用する方法であって、

仮想ターゲットアドレスで所定の仮想ターゲットに宛てられた、第１のプロトコルに従ってフォーマットされたパケットを入口側ラインカードにて受信するステップと、

前記入口側ラインカードが、フローＩＤを含む前記仮想ターゲットに関する情報を仮想ターゲット記述子から検索して、仮想ターゲット記述子識別子及び前記フローＩＤを前記パケットのローカルヘッダに付加するステップと、

前記入口側ラインカードが、前記フローＩＤに従って前記パケットを出口側ラインカードに送信するファブリックへ前記パケットを送信するステップと、

前記出口側ラインカードが、前記仮想ターゲット記述子識別子を使用して、前記仮想ターゲットに関連する物理的ターゲットが第２のプロトコルに従ってフォーマットされたパケットを必要とするか否かを含む、前記物理的ターゲットに関する情報を識別すると共に、前記物理的ターゲットに関する前記情報を使用して、仮想ターゲットブロックアドレスを物理的ターゲットブロックアドレスに変換し、必要であれば、前記パケットのフォーマットを前記第１のプロトコルから前記第２のプロトコルに変換するステップと、

前記出口側ラインカードが、前記物理的ターゲットブロックアドレスを使用して、前記パケットを前記物理的ターゲットに送信するステップと、

を含むことを特徴とする方法。

【請求項３５】

前記ステップの全ては、バッファリングを行うことなく実行されることを特徴とする請求項３４に記載の方法。

【請求項３６】

前記ステップの全ては、ワイヤ速度で実行されることを特徴とする請求項３４に

記載の方法。

【請求項３７】

ポートと、

前記ポートと通信を行い、バッファリングを行うことなくパケットを第１のプロトコルから第２のプロトコルに変換する変換器と、

を備えることを特徴とするラインカード。

【請求項３８】

前記変換器とは別個のＣＰＵを更に含むことを特徴とする請求項３７に記載のラインカード。

【請求項３９】

前記変換は、前記第１のプロトコルに従ってフォーマットされた場合の前記パケットからのフィールドを、前記第２のプロトコルに従ってフォーマットされた場合の前記パケットのフィールドにマッピングするステップを含むことを特徴とする請求項３７に記載のラインカード。

【請求項４０】

前記第１のプロトコルはｉＳＣＳＩであり、前記第２のプロトコルはファイバ・チャンネルであることを特徴とする請求項３７に記載のラインカード。

【請求項４１】

前記第１のプロトコルはファイバ・チャンネルであり、前記第１のプロトコルはｉＳＣＳＩであることを特徴とする請求項３７に記載のラインカード。

【請求項４２】

ポートと、

前記ポートと通信を行い、バッファリングを行うことなくパケットを第１のプロトコルから第２のプロトコルに変換するための手段と、

を備えることを特徴とするラインカード。

【請求項４３】

第１のプロトコルに従って作動する第１の装置に結合されることができる第１のポートを有する第１のラインカードと、

前記第１のラインカードと通信を行うファブリックと、

前記ファブリックと通信を行うと共に、第２のプロトコルに従って作動する第２の装置に結合されることができる第２のポートを有する第２のラインカードと、を備えるスイッチであって、

前記第２のラインカードは、プロトコル変換器を含み、前記プロトコル変換器は、前記第１のプロトコルに従ってフォーマットされたパケットを前記ファブリックから受信するために結合された入力部と、前記第２のポートに結合され、前記第２のプロトコルに従ってフォーマットされると共に前記第１のプロトコルに従ってフォーマットされた前記パケットに対応するパケットを生成する出力部とを有し、

前記第２のラインカードは、前記プロトコル変換器とは別個のＣＰＵを更に含むことを特徴とするスイッチ。

【請求項４４】

前記スイッチは、前記第１のポートにおいて前記第１のプロトコルに従ってフォーマットされたパケットを受信し、前記第２のポートにおいて前記第２のプロトコルに従ってフォーマットされた前記パケットをワイヤ速度で生成するように設計されることを特徴とする請求項４３に記載のスイッチ。

【請求項４５】

前記プロトコル変換器は、前記パケットのバッファリングを行うことなく作動することを特徴とする請求項４３に記載のスイッチ。

【請求項４６】

前記第１の装置はイニシエータであり、前記第２の装置はターゲットであることを特徴とする請求項４３に記載のスイッチ。

【請求項４７】

前記第１の装置はターゲットであり、前記第２の装置はイニシエータであることを特徴とする請求項４３に記載のスイッチ。

【請求項４８】

各々が少なくとも１つのポートを含む複数のラインカードと、

第１のラインカード上の第１のポートにおいて、第１のプロトコルに従ってフォーマットされた特定の仮想ターゲット宛のパケットを受信し、第２のラインカー

ド上の第２のポートにおいて、第２のプロトコルに従ってフォーマットされた対応するパケットを前記仮想ターゲットに関連する物理的ターゲットにワイヤ速度で送信するための手段と、
を備えることを特徴とするスイッチ。

【請求項４９】

プロセッサによって実行可能であり、データを記憶してアクセスするためのシステムに使用するスイッチ内の少なくとも１つの媒体上に記憶されたソフトウェア命令集合であって、
ｉＳＣＳＩ又はファイバ・チャンネルプロトコルの１つに従ってフォーマットされたパケットを第１の装置から受信するための命令と、
前記パケットが前記ｉＳＣＳＩプロトコルに従ってフォーマットされる場合、前記ｉＳＣＳＩプロトコルに従ってフォーマットされた前記パケットからのフィールドの少なくとも一部をファイバ・チャンネルパケットにマッピングすることによって、前記パケットを前記ファイバ・チャンネルパケットに変換するための命令と、
前記パケットが前記ファイバ・チャンネルプロトコルに従ってフォーマットされる場合、前記ファイバ・チャンネルプロトコルに従ってフォーマットされた前記パケットからのフィールドの少なくとも一部をｉＳＣＳＩパケットにマッピングすることによって、前記パケットを前記ｉＳＣＳＩパケットに変換するための命令と、
前記パケットを変換された状態で第２の装置に送信するための命令と、
を備え、前記命令の全ては、前記パケットのバッファリングを行うことを必要としないことを特徴とする命令。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、記憶領域ネットワーク（ＳＡＮ）に関する。

【背景技術】

【０００２】

集中的なデータ利用が急速に成長し、生データ記憶容量に対する需要が高まり続けている。企業が、電子商取引、オンライントランザクション処理、及びデータベースにますます依存するにつれて、管理及び記憶する必要がある情報量は莫大になる。その結果、記憶装置、より多くのユーザへのサービス提供、及び大量なデータのバックアップ作業を追加するための継続的な要求が作業を困難にするようになってきた。

【０００３】

データに対するこのような需要の増大に対応するために、記憶領域ネットワーク（ＳＡＮ）概念への人気が高まっている。ＳＡＮはストレージネットワーキング・インダストリ・アソシエーション（ＳＮＩＡ）によって定義されたネットワークであり、その主目的はコンピュータシステムと記憶素子との間、及び記憶素子間のデータ伝送である。例えば、ＳＣＳＩ接続による記憶装置とサーバとの直接的な接続、及びイーサネット（登録商標）（例えば、ＮＡＳシステム）等の従来型インタフェースによるＬＡＮへの記憶装置の増設とは異なり、ＳＡＮは、直接接続型のＳＣＳＩ及びＮＡＳと同様の帯域幅制限をもちにくい実質的に独立したネットワークを構築する。

【０００４】

詳細には、ＳＡＮ環境において、記憶装置（例えば、テープドライブ及びＲＡＩＤアレイ）及びサーバは、種々のスイッチ及び機器を介して相互接続されるのが一般的である。スイッチ及び機器との接続部は通常ファイバ・チャンネルである。一般的に、この構成により、ＳＡＮ上の任意のサーバは任意の記憶装置と通信することができ、その逆も同様である。また、サーバから記憶装置への代替経路を提供する。換言すると、特定のサーバが低速であるか又は完全に利用不可能

である場合、ＳＡＮ上の別のサーバが、記憶装置へのアクセスを提供することができる。また、ＳＡＮは、データのミラーリングができるようにし、利用可能な複数のコピーを作り、結果的にデータ利用における信頼性をより高める。より多くの記憶装置を必要とする場合、特定のサーバと接続する必要なく追加の記憶装置をＳＡＮに増設することができろ。それどころか、新しい記憶装置を記憶ネットワークに簡単に増設することができるとともに、任意の地点からアクセスすることができる。

【０００５】

図１の機能ブロック図のシステム１００にＳＡＮの一例を示す。図示したように、１又はそれ以上のサーバ１０２が存在する。例示的に３つのサーバ１０２のみが示されている。サーバ１０２は、イーサネット接続によってＬＡＮ１０６及び／又はルータ１０８に、そしてさらに、インターネット等のＷＡＮ１１０に接続される。さらに加えて、各々のサーバ１０２は、ファイバ・チャンネル接続によって、ＳＡＮの「編地（fabric）」と称されたりもする複数のファイバ・チャンネルスイッチ１１２の各々に接続される。例示的に２つのスイッチ１１２のみが示されている。次に、各々のスイッチ１１２は、複数のＳＡＮ機器１１４の各々に接続されている。例示的に２つの機器１１４のみが示されている。また、各々の機器は、テープドライブ、光学ドライブ、又はＲＡＩＤアレイ等の複数の記憶装置１１６の各々に結合される。更に、各々のスイッチ１１２及び機器１１４はゲートウェイ１１８に結合され、ゲートウェイ１１８はルータ１０８に結合され、最終的にルータ１０８はインターネット等の広域ネットワーク（ＷＡＮ）１１８に接続される。図１は、スイッチ１１２、機器１１４、記憶装置１１６、及びゲートウェイ１１８を含むＳＡＮ１１９として考えられる構成の一例を示す。なお、他の構成も可能である。例えば、１つの機器は、全スイッチ数よりも少ないスイッチに接続することができる。

【０００６】

機器１１４はＳＡＮの記憶管理を行う。機器１１４は、データを受信すると機器内のメモリにそのデータを記憶する。次に、正しい記憶装置にそのデータを転送するために、（機器内の）プロセッサを用いて当該データを分析し且つ操作す

る。一般的には、この記憶及び転送処理によりデータアクセスが減速してしまう。

【０００７】

機器はいくつかのスイッチングを行うが、多数のサーバ（３つ以上）が存在する場合があるため、及び、各々の機器のポート数が少ない（通常は２つ又は４つ）ため、スイッチ１１２は多数のサーバを少ない機器に接続する必要がある。それにもかかわらず、スイッチ１１２には殆ど処理能力（intelligence）が組み込まれておらず、選択された機器１１４にデータを転送するだけである。

【０００８】

機器が有する制限の１つは、一般に機器のポート数が非常に少ない（例えば、２つのポートのみ）という事実にある。その結果、機器が利用できる帯域幅は制限される場合がある。機器にポートを増設することは可能ではあるが、一般に非常に経費がかかるどの１又は２ポートも、高価なＣＰＵ又はサーバカードによってサポートされる。そこで一般に、ポートを増設するためには、（仮想化、並びに記憶及び転送機能を行う）全ファイルカードを装置に付加しなければならず、非常に経費がかかってしまうのが一般的である。もしくは、機器をＳＡＮに単に増設することもできるが、やはり非常に割高になりがちである。

【０００９】

更に、通常は機器１１４内において、ＳＡＮは「仮想化」として既知の機能を実行するのが一般である。１又はそれ以上の物理的記憶装置上の空間が特定のユーザに割り当てられた場合に仮想化が実行されるが、この空間の物理的位置はユーザには分からないままである。例えば、ユーザは、自社の「エンジニアリング記憶空間」ＥＮＧにアクセスする、即ち、ユーザが外付けディスクドライブにアクセス又は「参照」しようとする場合、仮想空間ＥＮＧにアクセス及び「参照」することができる。しかしながら、ＥＮＧ空間は、複数の物理的記憶装置上に分散させることができ、又は単一の記憶装置上に断片化させることさえできる。つまり、サーバが仮想装置（例えば、ＥＮＧ）及びブロック番号を要求する場合、機器は、要求された仮想装置と物理的に相関関係がある装置を特定して、それに応じてデータを導く必要がある。

【００１０】

一般に、ＳＡＮは装置を相互接続する単一のプロトコルを使用して構築される。ファイバ・チャンネルは最も一般的に使用されているが、同様にイーサネット接続も使用されている。しかしながら、両方のプロトコルの使用が望まれる場合、２つのプロトコル間である種の変換を行う必要がある。この場合、ファイバ・チャンネルＳＡＮ１１９は、一般にブリッジ１２１を経由してイーサネットＳＡＮ１２２に結合される。一方のプロトコルから他方のプロトコルに変換するために、ブリッジにより受信されたパケットはメモリに記憶される。パケットがメモリに一旦記憶されると、プロセッサは、パケットを操作して、一方のプロトコルのヘッダを取り除きそして他方のプロトコルのヘッダを組立て、この結果全く新しいパケットを作る。詳細には、図２を参照すると、ブリッジ１２１によって（１又はそれ以上のパケットから成る）要求が受信されると、例えば、ホストバスアダプタ（ＨＢＡ）２０２はこの要求をファイバ・チャンネル接続２０４上で受信する。プロセッサ２０８が要求を分析して操作する準備ができるまで、即ち、送信プロトコルに基づいて要求を再構築するまで、全ての要求がメモリ２０６に記憶される。プロセッサ２０８によって前記要求が操作されると、この要求はネットワークインタフェースカード（ＮＩＣ）２１０に送信され、その後にイーサネット接続２１２上に送出される。勿論、逆の場合（イーサネットからファイバ・チャンネルへ）も同じ処理を行うことができる。従って、プロトコル間の変換処理は、かなりのメモリ及びプロセッサリソースを必要とし、このことは、データ伝送の遅延を引き起こすだけでなく、金銭及び不動産の両面におけるシステムのコストアップにつながる。しかしながら、現在利用可能な唯一の選択肢は、プロトコルを別個の各ネットワーク上で孤立させ続けることである。

【００１１】

ゲートウェイ１１８（図１）は、ＳＡＮをＷＡＮに接続するだけでなく、２又はそれ以上のＳＡＮを相互に接続するために使用される場合が多い。通常、ゲートウェイは、種々のプロトコル変換を行うのではなく、むしろ、本技術分野では既知であるように、ＩＰパケットのデータをカプセル化する。それでもなお、複数のＳＡＮが接続される場合には、各々の接続装置に対して固有アドレスが存在す

る必要がある。しかしながら、ＩＰプロトコルはアドレス指定用に３２ビットを備えるが、ファイバ・チャンネルプロトコルは２４ビットのビット数のみを備えるにすぎない。従って、大半のＳＡＮはファイバ・チャンネルを使用することから、ゲートウェイの使用にもかかわらず拡張性が問題になり、インターネット上でのＳＡＮの使用を制限する場合がある。

【００１２】

ＳＡＮは数年前に導入されたが、普及に際して相互接続性の問題、利用可能な技術の不足、及び高い実行コストが大きな障害になっている。例えば、既存のＳＡＮは、配備コストが高く、管理コストも高い。図１を再度参照すると、一般的に、各々のスイッチ、機器、及びゲートウェイは、異なるベンダから提供され、管理基準の不足からベンダ専用の管理ツールが横行している。その結果、ＳＡＮを配備するには、複数のベンダから機器を購入する必要がある。また、図１に示すように、各々のスイッチ、機器、ゲートウェイ、記憶装置、サーバ、及びルータは、管理ステーション１２０として示されるような独自の管理をもつことになる。独立した物理的な管理ステーションが示されているが、この独立した管理は、単一のコンピュータ上の独立したベンタ専用ソフトウェアの形態である場合が多く、そのソフトウェアは相互に連絡を行わないことを理解されたい。その結果、ＳＡＮの集中管理が行われず、管理のために多くの人々を要求する複数の管理ステーションが存在するのが普通であれば、その管理コストが高くなってしまう。

【００１３】

（関連出願の説明）

本出願は、２００１年９月２８日出願の米国仮特許出願番号６０／３２５，７０４「記憶領域ネットワークのために記憶装置スイッチ」の優先権を主張するものであり、その開示内容は、引用により本明細書に組み込まれている。

また、本出願は、本出願と同時出願され、その開示内容が引用により本明細書に組み込まれている以下の出願に関連する。

米国出願番号１０／０５１，３２１「記憶領域ネットワークのための記憶装置スイッチ」

米国出願番号１０／０５１，１６４「サーバ不要の記憶サービス」

米国出願番号１０／０５１，０９３「記憶システムにおけるパケット分類」

米国出願番号１０／０５１，３９６「記憶システムにおける仮想化」

米国出願番号１０／０５１，３３９「記憶ネットワークにおけるサービス実行品質」

米国出願番号１０／０５０，９７４「記憶ネットワークにおける記憶リソースのプール及び準備」

米国出願番号１０／０５１，０５３「記憶リソースにおけるロードバランシング」

【００１４】

（発明の概要）

本発明の実施形態による記憶装置スイッチは、配備が容易で集中管理が可能なＳＡＮの構築を可能にする拡張性の高いスイッチである。更に、この記憶装置スイッチは、世界的規模のインフラ基盤の配備も可能にするので、記憶装置等のＳＡＮのリソースを実質的に世界中のどこにでも配置することが可能になる。更に、本発明による記憶装置スイッチにより、例えば、ｉＳＣＳＩ（イーサネット接続上で搬送される最も新しく導入されたプロトコル）又はファイバ・チャンネルの両方を含む多重プロトコルＳＡＮが、任意のデータパケットを「ワイヤ速度」で処理すること、即ち、単にスイッチング又は経路指定機能を実行するスイッチによって生ぜしめる待ち時間の増加を招くことなく処理を行うことが可能になり、本発明によるスイッチは高帯域幅を有する。一般的に、ワイヤ速度でデータを処理するために、本発明の実施形態による記憶装置スイッチは、従来とは異なり、パケットのバッファリングを行わない。従って、従来技術に比較して、本発明の実施形態によるアーキテクチャでは、パケットを処理するための所要時間が最小になる。

【００１５】

詳細には、本発明によるスイッチは、ワイヤ速度で仮想化及び変換サービスを行う。このようなワイヤ速度の処理を行うために、スイッチ・ラインカードの全ポートには「処理能力（intelligence）」が分散されている。更に、各々のライン

カードは、パケットを分類し且つデータパケットを制御パケットと分離することができる。また、処理能力（intelligence）が分散されるおかげで、各々のラインカードは、データパケット上で必要な場合には、（仮想アドレスを物理的アドレスへ変換する）仮想化及び（第１のプロトコルの受信プロトコルを第２のプロトコルの発信パケットへ変換する）プロトコル変換を行い、そしてユーザ又はサーバが仮想化又は変換の必要性を意識することなく又は関与することなく行うことができる。処理能力を分散させたので、従来のＣＰＵ又はサーバカードよりも安価な多数のラインカードを作ることができ、例えば、多数のポートに対応するための記憶装置スイッチの拡張が容易になる。

【００１６】

更に、各々のスイッチは、ミラーリング、低速リンク上でのミラーリング、スナップショット、仮想ターゲットクローン化（複製）、第三者コピー、定期的なスナップショット及びバックアップ、並びにリストア等のサーバ不要の記憶サービスを提供できる。スイッチがこのようなサービスに関する要求を受信すると、サーバ又は管理ステーション等の任意の他の装置の支援なしでこのサービスを実行することができる。

本発明を、以下の図面を参照しながら例示的な特定の実施形態に関して説明する。

【発明を実施するための最良の形態】

【００１７】

本発明による記憶装置スイッチを含むシステム３００を図３に示す。図示するように、このシステムは、既存のシステムに比べて非常に単純化されている。１つの実施形態において、システム３００は、複数のサーバ３０２を含む。例示的に、３つのサーバ３０２が示されているが、別の実施形態において更に多くの又は更に少ないサーバを使用することができる。また、図示されていないが、サーバはＬＡＮに結合することもできる。図示のように、各々のサーバ３０２は、記憶装置スイッチに接続されている。しかしながら、別の実施形態において、各々のサーバ３０２は、存在する記憶装置スイッチ３０４の全数よりも少ない数のスイッチに接続することができる。サーバとスイッチとの間に形成される接続は、

任意のプロトコルを利用することができるが、１つの実施形態において、接続は、ファイバ・チャンネル又はギガビットイーサネット（ｉＳＣＳＩプロトコルに従いパケットを伝送）のいずれかである。他の実施形態では、インテル社によって定義されたインフィニバンド・プロトコル、又は他のプロトコル若しくは接続を使用することができる。図示の実施形態において、各々のスイッチは、複数の記憶装置又はサブシステム３０６の各々に順次接続されている。しかし、他の実施形態において、各々のスイッチは、記憶装置又はサブシステム３０６の全数よりも少ない記憶装置又はサブシステムに接続することができる。記憶装置スイッチと記憶装置との間に形成される接続は、任意のプロトコルを利用することができるが、１つの実施形態において、接続は、ファイバ・チャンネル又はギガビットイーサネットのいずれかである。特定の実施形態において、１又はそれ以上のスイッチ３０４は、都市規模ネットワーク（ＭＡＮ）、又はインターネット３０８等の広域ネットワーク（ＷＡＮ）にそれぞれ結合されている。記憶装置スイッチとＷＡＮとの間で形成される接続は、大半の実施形態においてはインターネットプロトコル（ＩＰ）が使用されるのが一般的であろう。ＭＡＮ／ＷＡＮ３０８と直接接続されるように示されているが、他の実施形態では、スイッチ３０４とＭＡＮ／ＷＡＮ３０８との間の媒介手段としてルータ（図示せず）を利用することができる。更に、それぞれの管理ステーション３１０は、各々の記憶装置スイッチ３０４、各々のサーバ３０２、及び各々の記憶装置３０６に接続されている。管理ステーションは異なるコンピュータとして示されているが、各種装置を管理するためのソフトウェアは、まとめて１つのコンピュータ上に存在してもよいことを理解されたい。

【００１８】

図４は、本発明によるシステムの他の実施形態を示す。この実施形態において、２つのＳＡＮ４０２、４０４が形成されており、各々は、本発明の実施形態による１又はそれ以上の記憶装置スイッチ３０４を使用する。ＳＡＮ４０２及び４０４は、スイッチ３０４によってインターネット等のＷＡＮ３０８を経由して接続されている。接続は、任意の規格又はプロトコルとすることができるが、１つの実施形態において、パケット・オーバーＳＯＮＥＴ（ＰｏＳ）又は１０ギガビッ

トイーサネットである。

【００１９】

図５は、本発明によるシステムの更に別の実施形態を示し、スイッチ３０４は相互に直接結合されている。図３又は図４に示す実施形態のいずれかにおいて、１より多いスイッチが使用される場合、これらのスイッチは、図５に示すように結合することができる。

【００２０】

本発明による記憶装置スイッチは、世界的に分散された共有記憶プールとして使用可能な記憶装置の集中管理を可能にし、世界的に分散された莫大な量の管理ステーション及び大勢の熟練した管理要員を有する代わりとなる。このような記憶装置スイッチは、「処理能力のある（intelligence）」スイッチであり、図３を図１と比較すれば分かるように、スイッチ、機器、及びゲートウェイの機能は、本発明の実施形態による記憶装置スイッチ３０４内に有効に統合されている。このような記憶装置スイッチ３０４は、スイッチング機能に加えて、一般的に従来のアーキテクチャの機器によってもたらされる仮想化及び記憶サービス（例えば、ミラーリング）を提供すると共に、プロトコル変換も提供する。また、本発明の特定の実施形態による記憶装置スイッチは、付加的な機能（例えば、仮想プライベートネットワークによるデータセキュリティ）を実行する。このような付加的機能としては、サーバによって従来より行われている負荷バランス（load balancing）等の従来型システムにおける他の装置による機能、並びに従来型システムではこれまで利用不可能であった他の機能を挙げることができる。

【００２１】

本発明の実施形態による記憶装置スイッチの処理能力（intelligence）は、全てのスイッチポートに分散される。この分散された処理能力は、システムの拡張性及び利用可能性を認めている。

【００２２】

更に、処理能力を分散したので、本発明の実施形態によるスイッチは、「ワイヤ速度」でデータを処理することができる。このことは、記憶装置スイッチ３０４が、一般的なネットワークスイッチ（図１のスイッチ１１２等）により生ぜしめ

るのと同程度の待ち時間をデータパケットにもたらすことを意味する。つまり、スイッチの「ワイヤ速度」は、特定ポートへの接続によって評価される。したがって、ＯＣ－４８接続を有する１つの実施形態において、記憶装置スイッチは、ＯＣ－４８速度（２．５ビット／ナノ秒）に遅れずについていくことができる。ＯＣ－４８速度で送出される（１０ビット／バイトによる）２キロバイトのパケットがスイッチに到来する所要時間は僅か８マイクロ秒である。１キロバイトのパケットの所要時間は僅か４マイクロ秒である。１００バイトの最小パケットは、ほんの４００ナノ秒しか要しない。しかし、本明細書において用語「ワイヤ速度」処理が用いられる場合は、この処理が１００バイトのパケットを処理するのに僅か４００ナノ秒であることを意味していない。１つの実施形態において、記憶装置スイッチは、ＯＣ－４８速度、即ち約６マイクロ秒（４マイクロ秒／キロバイト又は２．５ビット／ナノ秒）で、（１バイトが１０ビットになるように１０ビット符号化した場合）１５００バイトの最大イーサネットパケット、を処理できることを意味する。処理が１ビット／ナノ秒として定義されることが一般的な１Ｇｂイーサネットポートを有する実施形態において、当該ポートの「ワイヤ速度」は１０マイクロ秒／キロバイトになり、これはスイッチがキロバイトを処理するために最大１０マイクロ秒であることを意味する。２Ｇｂファイバ・チャンネルポートを有する実施形態では、「ワイヤ速度」が５マイクロ秒／キロバイトになる。更に別の実施形態では、１０ギガビットイーサネット若しくはＯＣ－１９２速度、又はそれ以上でデータを処理することができる。

【００２３】

本明細書で用いる場合、「仮想化」は、ユーザが承認（subscribe）した仮想ターゲット空間を１又はそれ以上の物理的な記憶ターゲット装置上の空間へマッピングすることを実質的には意味する。「仮想」及び「仮想ターゲット」という用語は、承認毎に割り当てられた記憶空間が、記憶装置スイッチ３０４に接続する１又はそれ以上の物理的記憶ターゲット上のどこかに存在し得ることに由来する。物理的空間は、１又はそれ以上の「論理ユニット」（ＬＵ）を含み得る「仮想ターゲット」として提供することができる。各々の仮想ターゲットは、１又はそれ以上のＬＵ番号（ＬＵＮ）で特定される１又はそれ以上のＬＵから成り、ｉ

ＳＣＳＩプロトコル及びＦＣプロトコルで使用される場合が多い。各々の論理ユニット、従って各々の仮想ターゲットは、１又はそれ以上の領域、即ち物理的装置上の記憶空間の連続部分で構成されるのが一般的である。従って、仮想ターゲットは、記憶装置の全て（１つの領域）、単一の記憶装置の一部（１又はそれ以上の領域）、又は複数の記憶装置の一部（複数の領域）を占めることができる。物理的装置、ＬＵ、領域数、及びそれらの正確な位置は実体がなく、しかも承認ユーザに対して非表示である。

【００２４】

記憶空間は、多数の異なる物理的装置によってもたらされる場合があるが、各々の仮想ターゲットは、１又はそれ以上の領域（ドメイン）に属する。同じ領域のユーザのみが、その領域内の仮想ターゲットを共有することができる。領域集合によって、複数領域のユーザ管理が容易になる。領域集合に属するメンバは、同様に他の領域のメンバになることができる。しかし、本発明の実施形態において、仮想ターゲットは、１つの領域のみとすることができる。

【００２５】

図６は、本発明の実施形態による記憶装置スイッチ３０４の機能ブロック図を示す。１つの実施形態において、記憶装置スイッチ３０４は、複数のラインカード６０２、６０４、６０６、複数のファブリックカード６０８、及び２つのシステム制御カード６１０を含む。以下、それぞれについてその詳細を説明する。

【００２６】

システム制御カード：

２つのシステム制御カード（ＳＣＣ）６１０の各々は、全てのラインカード６０２、６０４、６０６に接続する。１つの実施形態において、このような接続は、本技術分野で公知のI^2C信号によって、ＳＣＣを備えたイーサネット接続によって形成される。ＳＣＣは、ファブリックカード（fabric card）と同様、I^2C接続によって電源投入を制御し、且つ個々のラインカードをモニタする。また、イーサネット接続上のカード間通信を用いて、ＳＣＣは、以下に詳細に説明するスナップショット及び複製等の種々の記憶サービスを開始する。

【００２７】

更に、ＳＣＣは、サーバや記憶装置等のようなスイッチに取り付けられた全ての仮想ターゲット及び物理的装置の構成情報を追跡するデータベース６１２を保持する。

【００２８】

更に、データベースは、仮想ターゲット及びユーザについての異なる領域及び領域集合に関する情報と同様、使用、エラー、及びアクセスデータに関する情報を保持する。データベースの記録を、本明細書では「オブジェクト」と呼ぶ。各々のイニシエータ（例えば、サーバ）及びターゲット（例えば、記憶装置）は、世界中で唯一の識別子（World Wide Identifier：WWUI）を有し、それは本技術分野では既知である。データベースはＳＣＣ内のメモリ素子に保持され、１つの実施形態においてメモリ素子はフラッシュメモリから形成されるが、他のメモリ素子でも十分である。

【００２９】

管理ステーション（３１０）は、イーサネット接続を用いてＳＣＣ６１０を介して記憶装置スイッチ３０４に達することができる。したがって、ＳＣＣは管理ステーションと接続するための付加的なイーサネットポートを含む。管理ステーションの管理者は、ＳＣＣデータベース６１２に記憶された任意のオブジェクトに対する事実上の照会且つ更新と同様、記憶装置又は仮想ターゲットの追加又は取り外しを知ることができる。

【００３０】

２つのＳＣＣ６１０のうち、一方は主作動ＳＣＣであるが他方はバックアップ用であり、記憶装置内において作動が同期し続けるが、作動を直接制御するようにはなっていない。ＳＣＣは、高い利用モードで作動し、一方のＳＣＣが故障した場合に他方は主制御装置になる。

【００３１】

ファブリックカード：

スイッチ３０４の１つの実施形態において、３つのファブリックカード６０８が存在するが、他の実施形態では、更に多くの又は更に少ないファブリックカードを有することができる。１つの実施形態において、各々のファブリックカード6

０８は、ラインカード６０２、６０４、６０６の各々に結合されており、ラインカードの全てを相互に接続する働きをする。１つの実施形態においてファブリックカード６０８は、全てのラインカードが存在する場合は、最大トラフィックをそれぞれ処理することができる。各ラインカードによって処理されるこのトラフィック負荷は、１つの実施形態において最大１６０Ｇｂｐｓであるが、他の実施形態では、これより多かったり少なかったりする最大トラフィック量を処理することができる。１つのファブリックカード６０８が故障すると、残りの２つのカードは、起こり得る最大スイッチトラフィックに対する十分な帯域幅をなおも有する。即ち、１つの実施形態において、各ラインカードは、入口側で１０Ｇｐｓ及び出口側で１０Ｇｐｓの２０Ｇｐｓのトラフィックを生成する。しかしながら、正常時には、３つのファブリックカードの全ては同時に有効である。各々のラインカードより、データに適合し得る３つのファブリックカードの任意の１つにこのデータトラフィックを送信する。

【００３２】

ラインカード：

ラインカードは、サーバ及び記憶装置への接続を形成する。１つの実施形態において、記憶装置スイッチ３０４は最大１６枚のラインカードをサポートするが、他の実施形態では、異なる数のラインカードをサポートすることができる。更に、１つの実施形態において、異なる３種類のラインカード、即ち、ギガビットイーサネット（ＧｉｇＥ）カード６０２、ファイバ・チャンネル（ＦＣ）カード６０４、及びWANカード６０４が利用される。他の実施形態では、これよりも多い又は少ない種類のラインカードを含むことができる。ＧｉｇＥカード６０２は、イーサネット接続用であり、１つの実施形態において、ｉＳＣＳＩサーバ又はｉＳＣＳＩ記憶装置（又は、他のイーサネット型装置）に接続する。ＦＣカード６０４は、ファイバ・チャンネル接続用であり、ファイバ・チャンネルプロトコル（ＦＣＰ）サーバ又はＦＣＰ記憶装置のいずれかに接続する。WANカード６０６は、MAN又はWANに接続するためのものである。

【００３３】

図７は、本発明による記憶装置スイッチ３０４の１つの実施形態において使用さ

れる汎用ラインカード７００の機能ブロック図を示す。この図は、ＧｉｇＥ６０２、ＦＣ６０４、又はWAN６０６等の全種類のラインカード間で共通の構成品を示す。他の実施形態において、他の種類のラインカードを利用して、インフィニバンド（Infiniband）等の他のプロトコルを用いて装置に接続することができる。各ラインカードの相違点を以下に説明する。

【００３４】

ポート：

各ラインカード７００は複数のポート７０２を含む。ポートはサーバ又は記憶装置のいずれかに対するラインカード接続を形成する。図示の実施形態では８つのポートが示されているが、他の実施形態では、これよりも多い又は少ないポートを使用することができる。例えば、１つの実施形態において、各々のＧｉｇＥカードは、最大８つまでの１Ｇｂイーサネットポートをサポートすることができ、各々のＦＣカードは、最大８つまでの１ＧｂのＦＣポート又は４つの２ＧｂのＦＣポートのいずれかをサポートすることができる。そして各々のWANカードは、最大４つまでのＯＣ－４８ポート又は２つのＯＣ－１９２ポートをサポートすることができる。従って、１つの実施形態において、接続可能な最大数は、スイッチ３０４毎に１２８ポートである。各々のラインカードのポートは全二重式であり、サーバ若しくは他のクライアントのいずれかに、又は記憶装置若しくはサブシステムに接続する。

【００３５】

更に、各々のポート７０２は関連メモリ７０３を有する。１つのポートに１つのメモリ素子だけが接続された状態で示されているが、各々のポートは、独自のメモリ素子を有することができること、又はポートの全ては、単一のメモリ素子に結合されてもよいことを理解されたい。説明を明瞭にする目的で、本明細書では１つのポートに１つのメモリ素子のみが接続された状態で示されている。

【００３６】

記憶処理ユニット：

１つの実施形態において、各々のポートは、記憶処理ユニット（ＳＰＵ）７０１につながっている。ＳＰＵは、データトラフィックを迅速に処理して、ワイヤ速

度の作動を可能にする。１つの実施形態において、ＳＰＵは、複数の構成要素、即ち、パケット統合及び分類エンジン（ＰＡＣＥ）７０４、パケット処理ユニット（ＰＰＵ）７０６、ＳＲＡＭ７０５、及びＣＡＭ７０７を含む。別の実施形態では、これよりも多い又は少ない構成要素を使用することができ、又は構成要素を組み合わせて同じ機能性を得ることができる。

【００３７】

ＰＡＣＥ：

各々のポートは、パケット統合及び分類エンジン（ＰＡＣＥ）７０４に結合されている。図示のように、ＰＡＣＥ７０４は、２つのポートを統合して２倍の帯域幅をもつ単一のデータチャンネルにする。例えば、ＰＡＣＥ７０４は、２つの１Ｇｂポートを統合して単一の２Ｇｂデータチャンネルにする。ＰＡＣＥは、以下に説明するように、受信した各々のパケットを制御パケット又はデータパケットに分類する。制御パケットは、ブリッジ７１６を経由してＣＰＵ７１４に送信され、処理される。データパケットは、ローカルヘッダが付加されて、以下に説明するパケット処理ユニット（ＰＰＵ）７０６に送信される。１つの実施形態において、ローカルヘッダは１６バイトであり、結果として、６４バイト（ヘッダ１６バイト、ペイロード４８バイト）のデータ「セル」又は「ローカルパケット」になる。ローカルヘッダは情報を伝えるために使用され、そしてスイッチ２０４により内部で使用される。ローカルヘッダは、パケットがスイッチを出る前に取り除かれる。従って、本明細書で用いる場合、「セル」又は「ローカルパケット」は、スイッチ内で局所的に使用される送信単位であり、ローカルヘッダ及び元のパケットを含む（特定の実施形態において、元のＴＣＰ／ＩＰヘッダも元のパケットから取り除かれる）。しかしながら、本発明の実施形態すべてが、ローカルヘッダを作成するか、又は外部パケットとは異なる「ローカルパケット」（セル）を有するものではない。したがって、本明細書で使用する場合、「パケット」という語は、「ローカル」パケット又は「外部」パケットを引用することができる。

【００３８】

分類機能は、スイッチが従来のシステムの記憶及び転送モデルを使用することな

く、記憶仮想化及びプロトコル変換機能をワイヤ速度で実行するのを助ける。各々のＰＡＣＥは、ＰＰＵ７０６への専用経路を有するが、図示の実施形態における４つのＰＡＣＥの全ては、ＣＰＵ７１４への経路を共有し、この経路は、１つの実施形態において１０４ＭＨｚ／３２（３．２Ｇｂｐｓ）ビットデータ経路である。

【００３９】

パケット処理ユニット（ＰＰＵ）：

ＰＰＵ７０６は、作動中（on-the-fly）に仮想化及びプロトコル変換を行うが、このことは、本処理に関してセル（ローカルパケット）がバッファリングされないことを意味する。また、後述するスイッチ型記憶サービス機能を実行する。ＰＰＵは、１つの実施形態において、ＯＣ－４８速度、即ち２．５Ｇｂｐｓで入口側及び出口側の両方向にセルを移動させることができるが、他の実施形態において、ＯＣ－１９２速度、即ち１０Ｇｂｐｓでセルを移動させることができる。１つの実施形態のＰＰＵは、入口側ＰＰＵ７０６$_1$及び出口側ＰＰＵ７０６$_2$を含み、両方を同時に実行する。入口側ＰＰＵ７０６$_1$は、ＰＡＣＥ７０４から入力データを受信してデータをトラフィックマネージャ７０８へ送信するが、出口側ＰＰＵ７０６$_2$は、トラフィックマネージャ７０８からデータを受信してデータをＰＡＣＥ７０４へ送信する。

【００４０】

多数の記憶装置の接続（例えば、サーバから仮想ターゲットへ）は、各々のポートで同時に確立できる。しかしながら、各々の接続は、仮想ターゲットに固有のものであり、（ｉＳＣＳＩ接続の場合）ＴＣＰ制御ブロックインデックス及びポート番号によって一意的に特定することができる。接続が確立されると、ラインカード７００のＣＰＵ７１４は、接続に関する仮想ターゲット記述子（ＶＴＤ）をＰＰＵ７０６に送信することによって有効な仮想ターゲットを知らせる。ＶＴＤは、ＰＰＵがデータ（例えば、仮想化、変換、及び種々の記憶サービス）上で適切に作動するのに必要な接続及び仮想ターゲットに関する全ての関連情報を含む。ＶＴＤは、ＳＣＣデータベース内のオブジェクトから得られ、かつＳＣＣデータベースの関連のオブジェクトに記憶されている情報の部分集合を含むのが普

通である。図７ａは、本発明の１つの実施形態におけるＶＴＤフィールドの一例を示す。しかしながら、本発明の他の実施形態は、これよりも多い又は少ないＶＴＤ、又はフィールドが異なるＶＴＤを有することができる。

【００４１】

ＶＴＤを記憶して素早くアクセスすることができるように、１つの実施形態において、ＰＰＵ７０６は、ＳＲＡＭ７０５及びＣＡＭ７０７に接続されている。ＳＲＡＭ７０５は、ＶＴＤデータベースを記憶する。また、ＶＴＤ識別子（ＶＴＤ　ＩＤ）のリスト、即ち又はアドレスは、ＶＴＤに素早くアクセスするためにＰＰＵ　ＣＡＭ７０７に保持される。ＶＴＤ　ＩＤは、ＴＣＰ制御ブロックのインデックス及びＬＵＮを使用して索引が付される（マッピングされる）。更に、ＩＰ経路指定サービスに関して、ＣＡＭ７０７は、ルートテーブルを含み、ＣＰＵはルートが追加又は取り除かれた場合にルートテーブルを更新する。

【００４２】

１つのＣＡＭ及びＳＲＡＭのみが１つのＰＰＵに接続されるように示されているが、これは説明を明瞭にするためであることに留意されたい。種々の実施形態において、各々のＰＰＵは、それぞれのＣＡＭ及びＳＲＡＭ装置に接続されることになるか、又はＰＰＵの全てが単一のＣＡＭ及び／又はＳＲＡＭに接続されることになる。

【００４３】

ＰＰＵに対する各未処理の要求（例えば、読み出し又は書込み）に関して、要求状態を追跡するためにタスク制御ブロックがＰＰＵ　ＳＲＡＭ７０７内に設けられる。入口側タスク制御ブロック（ＩＴＣＢ）は、入口側ＰＰＵ上の記憶装置スイッチが受信した要求状態を追跡し、出口側タスク制御ブロック（ＥＴＣＢ）は、出口側ＰＰＵ上の記憶装置スイッチが送信した要求状態を追跡する。各々の仮想ターゲット接続に関しては、多数の同時要求、従って、多数のタスク制御ブロックがあり得る。タスク制御ブロックは、要求開始時に割り当てられ、そして要求完了時に解放される。

【００４４】

トラフィックマネージャ：

各々のラインカード７００上には、２つのトラフィックマネージャ（ＴＭ）７０８があり、１つは入口側トラフィックＴＭであり、１つは出口側トラフィックＴＭである。１つの実施形態において、入口側ＴＭは、多重６４バイトデータセルの形態で、パケットを４つのＳＰＵの全てから受信する。この実施形態において、各々のデータセルは、１６バイトのローカルヘッダ及び４８バイトのペイロードを有する。ヘッダは、ＴＭにセルの宛先ポートを教えるフローＩＤを含む。また、特定の実施形態において、ＳＰＵは、セルをＴＭに転送する前にＴＭヘッダをセルに付与することができる。また、ＴＭ又はＳＰＵのいずれかは、特定の実施形態において、ファブリックカードを経由して送信できるようにセルを小さなセルに再分割することができる。

【００４５】

１つの実施形態において、入口側ＴＭは、１２８ビット、１０４Ｍｈｚインタフェース１７０を経由してデータセルをファブリックカードに送信する。出口側ＴＭは、データセルをファブリックカードから受信して、それらを４つのＳＰＵへ送る。

【００４６】

入口側ＴＭ及び出口側ＴＭの両方は、送出のためのセル（ローカルパケット）を待ち行列に入れるための大きなバッファ７１２を有する。入口側ＴＭ及び出口側ＴＭ用の両バッファ７１２は６４ＭＢであり、多数のパケットを待ち行列に入れることができる。通常、ＳＰＵは、ファブリックカードの送信フローが受信フローと同じ速さなので、セルを入口側ＴＭに素早く送信することができる。従って、セルは、出口側ＴＭに素早く移動する。これに対して、発信ポートが詰まったり、又は複数の入口側ラインカードから送られたりすることから、出口側ＴＭをバックアップすることができる。この場合、発信セルのヘッダにフラグを立てて、出口側ＳＰＵに迅速に対応策を取るように報知する。出口側ＴＭは、フロー制御機能の起動要求を入口側ＳＰＵへ送信する。インターネット上の通信トラフィックとは異なり、記憶トラフィックに関しては、パケット落ちは許容できないことに注目することは価値がある。従って、バッファ内のセル量が所定の閾値を超えると、ＳＰＵは、バッファオーバーフローを回避するために、受信トラフィ

ックを減速するようフロー制御機能を直ちに作動させる必要がある。

【００４７】

ファブリック接続：

ファブリック接続７１０は、TMの２５６ビットパラレル信号（入口側１２８ビット、出口側１２８ビット）を、バックプレーン（backplane）に対して１６ビットシリアルインタフェース（それぞれ入口側８ビット、出口側８ビット）に１６０Ｇｂｐｓで変換する。従って、バックプレーンは、１／１６のピンで実行中であるが速度は１６倍である。この変換によって、何千ものピンやワイヤを接続することなく、適切なコストで高い利用性のあるバックプレーンを構築することができる。更に、１つの実施形態では３つのファブリックカードが存在するので、１つの実施形態において、各々のラインカード上に３つの高速コネクタがあり、各々のコネクタは、８ビット信号を３つのファブリックカードのそれぞれ１つに接続する。勿論、別の実施形態では、３つのファブリック接続７１０を必要としない場合もある。

【００４８】

ＣＰＵ：

全てのラインカード上にはプロセッサ（ＣＰＵ）７１４があり、プロセッサは、１つの実施形態において、ＰｏｗｅｒＰＣ７５０Ｃｘｅである。１つの実施形態において、ＣＰＵ７１４は、バスコントローラ７１５及びブリッジ７１６を経由して、３．２Ｇｂバスで各々のＰＡＣＥに接続する。更に、ＣＰＵ７１４は、各々のＰＰＵ、ＣＡＭ、及びTMに接続するが、特定の実施形態において、この接続は、４０Ｍｂｐｓの低速である。３．２Ｇｂ及び４０Ｍｂ経路の両方によって、ＣＰＵは、ラインカード内の大半の素子と通信し、ラインカード上にある全素子の内部レジスタの読み出し及び書込みを行い、マイクロコードをダウンロードし、制御パケットを送受信することができる。

【００４９】

各々のラインカード上のＣＰＵは、電源投入時に全てのチップを初期化し、マイクロコードが必要であればマイクロコードをＳＰＵ及び各々のポートに責任をもってダウンロードする。ラインカードが実行状態になると、ＣＰＵは、制御トラ

フィックを処理する。仮想ターゲット接続を確立するために必要な情報に関して、ＣＰＵは、ＳＣＣから情報を要求し、次に、これはＳＣＣデータベース内の適切なオブジェクトから情報を取得する。

【００５０】

ラインカードとポートにおける区別：

１つの実施形態において、各種ラインカード内のポート、例えば、ＧｉｇＥ、ＦＣ、又はWANは、各々のラインカードが一種類のポートのみをサポートするように区別できる。１つの実施形態の各種ポートについて以下に説明する。勿論、他のラインカードポートを、他の実施形態のインフィニバンド等の他のプロトコルをサポートするように設計することもできる。

【００５１】

ＧｉｅＥポート：

ギガビットイーサネットポートは、ｉＳＣＳＩサーバ及び記憶装置に接続する。ＧｉｇＥポートは、全種類のイーサネットトラフィックを伝送するが、本発明の１つの実施形態による記憶装置スイッチ３０４によってワイヤ速度で一般に処理される唯一のネットワークトラフィックは、ＴＣＰ／ＩＰパケット内部のｉＳＣＳＩパケットデータユニット（ＰＤＵ）である。しかしながら、他の実施形態において、イーサネット接続で伝送される他のプロトコル（ネットワークファイルシステム（ＮＦＳ）等）によるパケットを、ＧｉｇＥポートで受信してＳＰＵ及び／又はＣＰＵで処理することができる。

【００５２】

ＧｉｇＥポートは、仮想ターゲット又はｉＳＣＳＩ装置についてのＴＣＰ／ＩＰセグメントを送受信する。仮想ターゲットに対するＴＣＰ接続を確立するために、ラインカードＣＰＵ７１４及びＳＣＣ６１０の両方が必要とされる。ＴＣＰパケットが受信され、初期接続手順が行われた後に、ＴＣＰ制御ブロックが作成され且つＧｉｇＥポートメモリ７０３に記憶される。また、接続を認証するとともに仮想ターゲットの構成を理解するために、ＳＣＣデータベースのオブジェクトからＶＴＤを検索して、ＣＰＴ　ＳＤＲＡＭ７０５に記憶する必要もある。ＴＣＰ制御ブロックは、パケットが属する特定のＴＣＰセッション又はｉＳＣＳＩ

接続を識別し、１つの実施形態において、ＴＣＰセグメント番号、状態、ウィンドウサイズ、及び場合によっては接続に関する別の情報を含む。更に、ＴＣＰ制御ブロックは、本明細書では「ＴＣＰ制御ブロックインデックス」として引用されるインデックスによって識別される。接続用ＶＴＤを作成して、ＳＰＵ　ＤＲＡＭ７０５に記憶する必要がある。ＣＰＵは、ＳＤＲＡＭに記憶され、当初はＳＣＣデータベースから取得したＶＴＤ情報を検索することによってＶＴＤを作成する。ＶＴＤ　ＩＤは、ＶＴＤを素早く参照するためにＳＰＵ　ＣＡＭ７０７内のＶＴＤ　ＩＤリストに設定される。ＶＴＤ　ＩＤは、ＴＣＰ制御ブロックインデックスに関連付けされ、かつＴＣＰ制御ブロックインデックスによってインデックスが付けられる。

【００５３】

ポートはｉＳＣＳＩ　ＰＤＵを受信すると実質的に接続の終端ポイントとして機能するが、その後、スイッチは、そのターゲットを用いて新しい接続を初期化する。入口側でパケットを受信した後に、ポートは、ｉＳＣＳＩ　ＰＤＵをＴＣＰ制御ブロックインデックスと共にＰＡＣＥに送出して、特定のＴＣＰ接続を識別する。非ＴＣＰパケット又はｉＳＣＳＩ　ＰＤＵを含まないＴＣＰパケットに対しては、ポートは、接続の終端ポイントとしての機能を果たすことなく、パケットを送受信する。一般に、ポート７０２はＰＡＣＥ７０４と通信し、ｉＳＣＳＩパケットは、ＴＣＰ制御ブロックインデックスを使用して送受信される。パケットのＴＣＰ制御ブロックインデックスが－１の場合は、非ｉＳＣＳＩパケットを特定する。

【００５４】

ＦＣポート：

ＦＣポートは、サーバ及びＦＣ記憶装置に接続する。ＦＣポートは、接続サーバにはファイバ・チャンネル記憶サブシステムとして見えるが、このことは、本技術分野では理解されているように、イニシエータが、接続を確立するためのプロセスログイン（ＰＬＯＧＩ又はＰＲＬＩ）が実行可能な仮想ターゲット装置の巨大プールのように見えることを意味する。ＦＣポートは、ＧＩＤ拡張リンクサービス（ＥＬＳ）を受け取り、そしてイニシエータ（例えば、サーバ）によるアク

セス可能なターゲット装置のリストを返す。

【００５５】

ファイバ・チャンネル記憶装置に接続すると、ポートは、ファイバ・チャンネルＦポートとして見えるが、このことは、本技術分野では理解されているように、記憶装置からファブリックログインを受け入れ、ＧＩＤ要求を受け入れ且つ処理することによってネームサービス機能を提供することを意味する。

【００５６】

ポート初期化時に、ラインカードＣＰＵは、ファブリックログイン、プロセスログイン、及びＧＩＤの送受信を行う必要がある。ＳＣＣは、ＦＣ　ＥＬＳをｉＳＣＳＩ要求及びｉＳＣＳＩ応答に変換するためのアプリケーションをサポートする。その結果、ＳＣＣ内の同じデータベースは、ｉＳＣＳＩイニシエータ及びターゲットであるかのように、ＦＣイニシエータ（例えば、サーバ）及びターゲット（例えば、記憶装置）を追跡し続ける。

【００５７】

ＦＣ接続が確立すると、ＦＣポートは、ＧｉｇＥポートとは異なり、ＴＣＰ制御ブロック又はその等価物を作成する必要はない。全ての必要な情報は、ＦＣヘッダから入手できる。しかし、（Ｄ＿ＩＤによりインデックスされた）ＶＴＤは、依然として、ＧｉｇＥポートに関して説明したのと同様の方法で確立する必要があるであろう。

【００５８】

ＦＣポートは、１Ｇｂポート又は２Ｇｂポートとして構成することができる。１Ｇｂポートとしては、図７に示すように、２つのポートが単一のＰＡＣＥに接続されているが、２Ｇｂポートとして構成される実施形態において、ポートトラフィック及びＳＰＵが対応可能なトラフィックは、ＳＰＵでの混雑を避けるために一致させるべきである。１つの実施形態において、ポートは、ＰＯＳ／ＰＨＹインタフェースでＰＡＣＥに接続する。各々のポートは、別個に構成することができ、即ち、１つのＰＡＣＥは、２つの１Ｇｂポートを有することができ、別のＰＡＣＥは、単一の２Ｇｂポートを有することができる。

【００５９】

WANポート：

WANラインカードを含む実施形態において、１つの実施形態では、WANラインカードは、ＯＣ－４８及びＯＣ－１９２接続をサポートする。従って、２種類のWANポート、即ち、ＯＣ－４８とＯＣ－１９２とが存在する。ＯＣ－４８に関して、各々のＳＰＵに対して１つのポートがある。ＰＡＣＥには統合機能がないが、依然として分類機能はある。WANポートは、ＳＯＮＥＴに接続し、そしてＩＣＭＰ、ＲＩＰ、ＢＰＧ、ＩＰ、及びＴＣＰ等のネットワークパケットの送受信時にＧｉｇＥポートのように機能する。ＧｉｇＥポートとは異なり、１つの実施形態におけるWANポートは、追加のハードウェア構成部品を必要とするＶＰＮ及びＩＰＳｅｓに対するネットワークセキュリティをサポートする。

【００６０】

ＯＣ－１９２のワイヤ速度は高速になるので、ＯＣ－１９２をサポートする実施形態においては高速ＳＰＵが必要となるであろう。

【００６１】

スイッチ型の記憶操作：

本発明の１つの実施形態による記憶装置スイッチは、パケットの分類、仮想化、及び変換を含め、様々なスイッチ型の記憶操作を行なう。これらのサービスは、一般的にＳＰＵによって行われる。１つの実施形態において、全てのポートはＳＰＵを有し、制御トラフィックを処理するためのリソースを有するＣＰＵへ制御トラフィックを送る間に、データトラフィックをできるだけ高速で処理することが可能になる。図７に示すように、４つのＳＰＵは、８つのポートをサポートする単一のＣＰＵを共有する。従って、データトラフィックは最小のリソース及びオーバーヘッドを使用し、各々がワイヤ速度で記憶トラフィックを処理するための処理能力を有する多数の低コストポートが可能になる。ＳＰＵ機能については以下で詳細に説明する。

【００６２】

ＳＰＵ機能を説明する前に、ｉＳＣＳＩ　ＰＤＵ（パケットデータユニット）及びＦＣ　ＩＵ（情報ユニット）の概要を説明することは好都合であろう。しかしながら、ｉＳＣＳＩ及びＦＣプロトコルに関する一般知識をもっていることが前

提である。ｉＳＣＳＩに関する詳細は、引用により本明細書に組み込まれている、インターネットエンジニアリングタスクフォース（ＩＥＴＦ）による続刊中のインターネットドラフト「ｄｒｆｔ－ｉｅｔｆ－ｉｐｓ－ｉＳＣＳＩ－０７．ｔｘｔ」、２００１年７月２０日、を参照されたい。ファイバ・チャンネル（ＦＣ）に関する詳細は、引用により本明細書に組み込まれている、「情報システム－ＳＣＳＩ用ｄｐＡＮＳファイバ・チャンネルプロトコル」改訂０１２、１９９５年１２月４日（米国規格協会提案のドラフト）を参照されたい。

関連のＰＤＵ及びＩＵを、以下簡単に説明する。

【００６３】

ｉＳＣＳＩコマンドＰＤＵ：

図８ａはｉＳＣＳＩコマンドＰＤＵを示す。図示のように、以下のフィールドを有する４８バイトを含む。第１のバイト（バイト０）において、Ｘビットがイニシエータからターゲットまでの ＰＤＵについての再試行／再起動インジケータとして使用される。Ｉビットは、即時送出マーカとして使用される。Ｏｐｃｏｄｅ　０ｘ０１は、ｉＳＣＳＩ　ＰＤＵの種類がコマンドであることを示す。また、バイト１は、複数のフラグ、Ｆ（最終）、Ｒ（読み出し）、及びＷ（書込み）を有する。また、バイト１は、通常３ビットであるタスク属性フィールドＡＴＴＲを有する。バイト３のＣＲＮは、ＳＣＳＩコマンド参照番号である。ＡＨＳ全長は、４バイト語における任意の追加の随意的なヘッダセグメント（図示せず）の全長を表す。データセグメント長は、ペイロードの長さを示す。ＬＵＮは、論理ユニット番号を指定する。イニシエータタスクタグは、タスクを識別するためにイニシエータ（例えば、記憶装置）によって割り当てられたタスクタグを識別する。期待データ伝送長は、操作に関するイニシエータに伝送されるか、又はそこから伝送されるデータのバイト数を示す。ＥｘｐＳｔａｔＳＮは期待状態シーケンス番号であり、そしてＥｘｐＤａｔａＳＮは、期待データシーケンス番号である。コマンド記述子ブロック（ＣＤＢ）は、一般に１６バイトであり、ＳＣＳＩコマンド自体を具体化する。

【００６４】

ｉＳＣＳＩ　Ｒ２Ｔ　ＰＤＵ：

図８ｂはｉＳＣＳＩ　Ｒ２Ｔ　ＰＤＵを示す。バイト０において、０ｘ３１は、パケットをＲ２Ｔパケットとして識別する。イニシエータタスクタグは、コマンドＰＤＵの場合と同じある。ターゲット伝送タグは、ターゲット（例えば、記憶装置）によって割り当てられ、データパケットの識別を可能にする。ＳｔａｔＳＮフィールドは、状態シーケンス番号を含む。ＥｘｐＣｍｄＳＮは、次の期待ＣｍｄＳＮをイニシエータから識別し、ＭａｘＣｍｄＳＮは、イニシエータから受け入れ可能な最大ＣｍｄＳＮを識別する。Ｒ２ＴＳＮは、Ｒ２Ｔ　ＰＤＵ番号を識別する。期待データ伝送長は、ターゲットがイニシエータに送信してもらいたいバイト数を指定する（ターゲットは、複数のかたまり（chunk）でデータを要求できる）。従って、ターゲットはデータ伝送を開始すべきポイントを示すバッファオフセットも指定する。

【００６５】

ｉＳＣＳＩデータ書込み及び読み出しＰＤＵ：

図８ｃはｉＳＣＳＩ書込みデータＰＤＵを示す。図８ｄはｉＳＣＳＩ読み出しデータＰＤＵを示す。０バイトにおいて、０ｘ０５は、パケットを書込みパケットとして識別し、０ｘ２５は、パケットを読み出しパケットとして識別する。これらのＰＤＵのフィールドの大半は、前述のＰＤＵの場合と同じである。更に、ＤａｔａＳＮは、データシーケンス番号を識別し、そして残余カウントは例えば、イニシエータの期待データ伝送長が短すぎた場合に、期待されたバイトの何バイトが伝送されなかったかを識別する。

【００６６】

ｉＳＣＳＩ応答ＰＤＵ：

図８ｅはｉＳＣＳＩ応答ＰＤＵを示す。バイト０において、０ｘ２１は、パケットを応答パケットとして識別する。状態フィールドは、コマンドのＳＣＳＩ状態を報告するために使用される。応答フィールドは、コマンドが完了したか、又は、エラー又は故障があったかを識別するｉＳＣＳＩサービス応答コードを含む。基本の残余カウントは、例えば、イニシエータの期待データ伝送長が短すぎた場合に、期待されたバイトの何バイトが伝送されなかったかを識別する。Ｂｉｄｉ＿読み出し残余カウントは、期待されたバイトの何バイトがイニシエータに

伝送されなかったか示す。他のフィールドは、前述の他のＰＤＵと同じである。

【００６７】

ＦＣＰフレームヘッダ：

各々のＦＣＰ情報ユニット（ＩＵ）は図８ｆに示すフレームヘッダを使用し、以下に説明するペイロードが続く。Ｒ＿ＣＴＬは、フレームをＦＣ操作の一部として識別し、情報カテゴリーを識別する。Ｄ＿ＩＤは、フレームの宛先を識別する。Ｓ＿ＩＤは、フレームのソースを識別する。ＴＹＰＥは、一般に、ＳＣＳＩ　ＦＣＰシーケンスの全フレームについては０ｘ８０に設定される。Ｆ＿ＣＴＬは、シーケンス及び交換の開始、及び正常終了又は異常終了を管理する。ＳＥＱ＿ＩＤは、特定の交換発信者と交換応答者との間の各々のシーケンスを固有値で識別する。ＤＦ＿ＣＴＬは、存在できる任意の随意的なヘッダを示す。ＳＥＱ＿ＣＮＴは、シーケンス内のフレーム順序を示す。ＯＸ＿ＩＤフィールドは、交換発信者（イニシエータ）識別子である。ＲＸ＿ＩＤフィールドは、交換応答者（ターゲット）識別子である。ＲＬＴＶ＿ＯＦＦフィールドは、情報カテゴリーの基本アドレスを基準にした各々のフレームのペイロードの最初のバイトの相対変位を示す。

【００６８】

ＦＣＰ　ＣＭＮＤペイロード：

図８ｇはＦＣＰコマンドＩＵのペイロードを示す。ＦＣＰ＿ＬＵＮは、論理ユニット番号である。ＦＣＰ＿ＣＮＴＬは、複数の制御フラグ及びビットを含む制御フィールドである。ＦＣＰ＿ＣＤＢは、アドレス指定された論理ユニットによって解釈されることになる実際のＳＣＳＩ　ＣＤＢを含む。ＦＣＰ＿ＤＬは、ターゲットに伝送されるか又はそこから伝送されることが期待されるデータバイトの最大数のカウントを含む。

【００６９】

ＦＣＰ　ＸＦＲ　ＲＤＹペイロード：

図８ｈはＦＣＰ　ＸＦＲ＿ＲＤＹのペイロードを示す。ＤＡＴＡ＿ＲＯフィールドは、次のＦＣＰ＿ＤＡＴＡ　ＩＵの最初のデータバイトのＲＬＴＶ＿ＯＦＦフィールドの内容を示す。ＢＵＲＳＴ＿ＬＥＮフィールドは、次のＦＣＰ＿ＤＡＴ

Ａ　ＩＵのために準備されたバッファ空間の大きさを示し、正確な長さのＩＵの伝送を要求する。

ＰＣＰ　ＤＡＴＡ　ＩＵ：

データＩＵに関するペイロードは、伝送された実際のデータである。

【００７０】

ＦＣＰ　ＲＳＰ　ＩＵ：

図８ｉはＦＣＰ応答ＩＵのペイロードを示す。ＦＣＰ＿ＳＴＡＴＵＳフィールドは、コマンドタスクの正常終了時で０に設定される。正常に終了しなかった場合には、種々のステータス状況を示す。ＦＣＰ＿ＲＥＳＩＤフィールドは、このＳＣＳＩコマンドに対してＦＣＰ＿ＤＡＴＡ　ＩＵにおいて伝送されなかった残余データバイト数のカウントを含む。ＦＣＰ　ＳＮＳ＿ＬＥＮは、ＦＣＰ＿ＳＮＳ＿ＩＮＦＯフィールドのバイト数を指定する。ＦＣＰ＿ＲＳＰ＿ＬＥＮは、ＦＣＰ＿ＲＳＰ＿ＩＮＦＯフィールドのバイト数を指定する。ＦＣＰ＿ＲＳＰ＿ＩＮＦＯフィールドは、検出された任意のプロトコル障害を記述する情報を含む。ＦＣＰ＿ＳＮＳ＿ＩＮＦＯフィールドは、任意のセンスデータの存在を含む。

【００７１】

各々のｉＳＣＳＩ　ＰＤＵ及びＦＣ　ＩＵの詳細は、概括的に説明されている。ｉＳＣＳＩ　ＰＤＵ、ＦＣ　ＩＵ、及びこれらの各々のフィールドに関する更に詳細な説明は、前述のｉＳＣＳＩ文献及びＦＣ文献に見出すことができる。

【００７２】

記憶装置スイッチの分類：

パケット又はフレーム（本明細書では全体として「パケット」と呼ぶ）が記憶装置スイッチに達すると、各々のポートにおいてデータ及び制御トラフィックに分離される。データトラフィックは、ワイヤ速度での仮想化及び変換のためにＰＰＵに経路指定されるが、接続要求又は記憶管理要求等のデータトラフィックは、ＣＰＵに経路指定される。この分離を本明細書では「パケット分類」又は単に「分類」と呼び、一般にＳＰＵのＰＡＣＥで初期化される。従って、全パケットをＣＰＵに送って処理する既存の技術とは異なり、本発明によるシステムは、データトラフィックを個別に高速で処理できるように、パケットコンテンツを認識

して、ワイヤ速度での処理を可能にすることを助長するようになっている。ＧｉｇＥパケット及びＦＣフレームは、以下に説明するように、若干異なる方法で処理される。

【００７３】

ＧｉｇＥポートの入口側に到達するパケット（スイッチに到達するパケット）に関して、図９ａを参照しながら以下で各ステップを説明する。１つの実施形態において、ＧｉｇＥポートは、ＩＰパケット又はｉＳＣＳＩパケットのいずれかであるパケットを受信する（ステップ９０２）。パケットを受信すると、ＰＡＣＥは、そのパケットとともに有効なＴＣＰ制御ブロックインデックス（例えば、－１ではないインデックス）をポートから受信したか否かによって、仮想ターゲットアクセスが認識されたか否かを判定する（ステップ９０４）。有効なＴＣＰ制御ブロックインデックスがある場合、次に、ＰＡＣＥは、パケットのＴＣＰヘッダのフラグをチェックする（ステップ９０６）。ＴＣＰヘッダのＳＹＮ、ＦＩＮ、及びＲＳＴフラグが設定されている場合、ＣＰＵはＴＣＰセッションの確立及び終了を行う責任があるので、パケットはＣＰＵに送られる（ステップ９１６）。ｉＳＣＳＩ　ＴＣＰセッションが確立されると、ＴＣＰセッションを管理するために、ＧｉｇＥポートは、有効なＴＣＰ制御ブロックをＣＰＵから受信することになる。しかし、フラグが設定されていない場合、１つの実施形態において、ＰＡＣＥは、ＴＣＰ、ＩＰ、及びＭＡＣヘッダを取り除き（ステップ９０８）、ｉＳＣＳＩヘッダを残し、次に、ローカルヘッダを追加する（ステップ９１０）。しかしながら、他の実施形態では、ＴＣＰ、ＩＰ、及びＭＡＣヘッダを残して単にローカルヘッダを追加することができる。ローカルヘッダが追加されると、パケットはＰＰＵに送信される（ステップ９１２）。

【００７４】

更に図１０ａを参照すると、ステップ９１０が実行されると、受信されたＴＣＰパケット１００２は、ローカルパケット１００４に変換され、ＩＰ、ＴＣＰ、及びＭＡＣヘッダ１００６、１００８、１００９が取り除かれ（１つの実施形態において）、ローカルヘッダ１０１０が追加される。しかしながら、場合によっては、ｉＳＣＳＩパケットに関するペイロードは、２つのＴＣＰ／ＩＰパケットに

分割することもできる。従って、図１０ｂを参照すると、受信されたＴＣＰパケット１０１２は、ペイロードの第２の部分１０１４を含む場合もあり、ペイロードの第１の部分は先行パケットで送信済みである。ペイロードの第２の部分を含むパケットは、独立した新たなペイロード１０１６を追加的に含む。受信されたパケット１０１２は、２つのローカルパケット１０１８及び１０２０に分割される。ローカルパケット１０１８は、ローカルヘッダ１０２２及び先行パケットからのペイロード１０２４の第２の部分を含むが、ｉＳＣＳＩヘッダは含まない。ローカルパケット１０２０は、ローカルヘッダ１０２６、ｉＳＣＳＩヘッダ１０２８、及び新しいペイロード１０３０を含む。

【００７５】

図１１は１つの実施形態において使用されるローカルヘッダ１１００の一例を示す。ローカルヘッダ１１００は、１つの実施形態において、以下のフィールドを含む。ＶＴＤ　ＩＤフィールドは、特定の接続に関するＶＴＤを識別するために使用される。フローＩＤは、パケットの宛先ポートを指定する。ＴＣＰ制御ブロックインデックスは、特定の接続に関するＴＣＰ制御ブロックを指定する（ＴＣＰ接続の場合）。ＴＹＰＥフィールドは、データ又は制御といったパケット分類を指定する。サイズフィールドは、パケットサイズを示す。タスクインデックスは、スイッチ内のパケットを追跡して方向付けするために、並びに特定のタスクのパケットに関連する記憶情報の位置を見つけるために使用される。ローカルヘッダは、更に、ソース識別子（例えば、ソースポート、ＰＡＣＥ、ラインカード、及び／又はＣＰＵを識別する）及び宛先識別子（例えば、宛先ポート、ＰＡＣＥラインカード、及び／又はＣＰＵを識別する）等の特定のハードウェア識別子を含む。

【００７６】

ローカルヘッダは、スイッチの全体にわたって、種々の装置（例えば、ＰＡＣＥ、ＰＰＵ）によって使用される。従って、ローカルヘッダの一部のフィールドを完全に使用できる場合もあり、フィールドコンテンツを交換又は更新できる場合もある。

【００７７】

再度図９ａを参照すると、有効なＴＣＰ制御ブロックインデックスがない場合（ステップ９０４）、パケットがＩＰパケットであるか否かが判定される（ステップ９１４）。パケットがＩＰパケットでない場合、ＣＰＵに送られる（ステップ９１６）。パケットがＩＰパケットである場合、次に、ＰＡＣＥは、宛先ＩＰアドレスをチェックする（ステップ９１８）。ＩＰアドレスが記憶装置スイッチのポートのＩＰアドレスと一致した場合、パケットは、ＣＰＵに送信され（ステップ９１６）、処理される。ＩＰアドレスが記憶装置スイッチのポートのＩＰアドレスと一致しない場合、これは経路指定トラフィックであり、ＰＰＵに送られる（ステップ９１２）。

【００７８】

図９ｂを参照すると、ＧｉｇＥポート宛のパケットが出口側でＰＡＣＥによってＰＰＵ又はＣＰＵから受信されると（ステップ９５０）、ＰＡＣＥは、ローカルヘッダを取り除く（ステップ９５２）。パケットがＴＣＰセッション用である場合（ステップ９５４）、ＰＡＣＥは、ＧｉｇＥポートにその旨を知らせるために、ポートとのインタフェースの制御フラグを設定する（ステップ９５６）。パケットがＴＣＰセッション用である場合、ＰＡＣＥは、インタフェース制御信号を使用してパケット及びＴＣＰ制御ブロックインデックスをポートに送る（ステップ９５８）。ＴＣＰセッションがない場合、パケットは単純にポートに送られる（ステップ９６０）。

【００７９】

図１２ａは、ＦＣポートから到着したパケットを分類する際にＰＡＣＥにて行われるステップを示す。ＧｉｇＥポートの場合とは異なり、ＦＣポートのＰＡＣＥは、ＴＣＰ制御ブロックインデックスを処理する必要はない。その代わりに、ＦＣポートにてパケットを受信すると（ステップ１２０２）、ＦＣＰフレームヘッダのＳ＿ＩＤフィールドは、フレームがオープンＦＣ接続に属するか否かを判定するために参照されることができるが、このステップは、パケットがＰＰＵに送られた後に実行される。従って、ＰＡＣＥは、フレームがＦＣＰフレームであるか否かだけを判定すればよく（ステップ１２０４）、このことは、フレームヘッダのＲ＿ＣＴＬ及びＴＹＰＥフィールドを参照することで判定することができ

る。ローカルヘッダ１１００（図１１）が追加されるが（ステップ１２０６）、ヘッダ内のデータはＰＰＵにとっては後で有用なので、ＦＣＰフレームヘッダはこの時点では取り除かない。次に、ローカルパケットはＰＰＵに送られる（ステップ１２０８）。ＦＣＰフレームではない場合、フレームはＣＰＵに送られる（ステップ１２１０）。

【００８０】

図１２ｂを参照すると、ＦＣポート宛のパケットが出口側でＰＡＣＥによってＰＰＵ又はＣＰＵから受信されると（ステップ１２５０）、ＰＡＣＥは、フレームをＦＣポートへ送る前に（ステップ１２５６）、単純にローカルヘッダを取り除く（ステップ１２５４）。ローカルヘッダは、ＰＡＣＥに対してパケットが（ＰＡＣＥが接続されている２つのポートのうちの）どのポート宛になっているのかを指示することになる。

【００８１】

ＧｉｇＥポート又はＦＣポートのいずれかで受信され、ＰＰＵに送られるパケットに関して、１つの実施形態において、ＰＰＵは、更に制御トラフィックを分離する。図１３ａを参照すると、ＰＰＵがパケットをＰＡＣＥから受信すると（ステップ１３０２）、ＰＰＵは、ＩＰパケットか又はＴＣＰパケットかを判定する（ステップ１３０４）。パケットがＩＰパケットである場合、ＰＰＵは、ＣＡＭを検索してルートテーブルからパケットのフローＩＤを取得する（ステップ１３０６）。検索に失敗した場合、パケットは、未知の宛先ＩＰアドレスを有し、これはＣＰＵに送られ（ステップ１３０８）、ＣＰＵは、ＩＣＭＰパケットをソースＩＰアドレスに逆送信する（ステップ１３１０）。検索によってフローＩＤが戻されると、パケットは、トラフィックマネージャに送られる（ステップ１３１１）。

【００８２】

受信されたパケットがＴＣＰパケットである場合（ステップ１３０４）、ＰＰＵはＴＣＰ制御ブロックインデックスを使用してＣＡＭを検索し、ＴＣＰ制御ブロックインデックスは、ＴＣＰセッションを識別し、ｉＳＣＳＩヘッダからのＬＵＮと一緒に、ＴＣＰセッションは、仮想ターゲットを識別して仮想ターゲット記

述子ＩＤ（ＶＴＤ　ＩＤ）を取得する（ステップ１３１２）。ＶＴＤ　ＩＤは、実質的に、ＰＰＵ　ＳＲＡＭに記憶されたＶＴＤをアドレス指定するか又は指示する。ＰＰＵは、ＶＴＤ　ＩＤを使用して、ＶＴＤのアドレスを取得するので（ステップ１３１２）、ＶＴＤ　ＩＤの検索によって素早くＶＴＤの位置を見つけることができる。ＶＴＤを取得できない場合にはｉＳＣＳＩセッションがまだ確立されておらず、パケットはＣＰＵに送られる（ステップ１３１４）。しかし、ＶＴＤ　ＩＤがステップ１３１２で取得された場合、ＰＰＵは、パケットがｉＳＣＳＩ　ＰＤＵを含むか否かを判定する（ステップ１３１５）。パケットがｉＳＣＳＩ　ＰＤＵを含んでいない場合、パケットはＣＰＵに送られる（ステップ１３１４）。しかし、パケットがｉＳＣＳＩ　ＰＤＵを含む場合、ＰＰＵは、ＰＤＵがＰＤＵを移動させるデータ（例えば、読み出しコマンド、書込みコマンド、Ｒ２Ｔ、書込みデータ、読み出しデータ、応答）であるか否かを判定する（ステップ１３１６）。ＰＤＵがＰＤＵ移動データでない場合、パケットは、ＣＰＵに送られる（ステップ１３１４）。しかし、ＰＤＵがＰＤＵ移動データである場合、以下に説明するように、ＰＰＵは、パケットに別の処理、例えば、仮想化及び変換処理を行なう（ステップ１３１８）。

【００８３】

ＰＰＵがＦＣＰコマンドＩＵを有するＦＣＰフレームを入口側で受信した場合、ＰＰＵは、図１３ａで説明するステップと同様のステップを実行する。即ち、ステップ１３１２においてＣＡＭ検索がＶＴＤ　ＩＤを見つけるためにＦＣＰフレームからのＳ＿ＩＤアドレス及びＬＵＮを使用する点を除く、ステップ１３０２、１３１２～１３１８である。

【００８４】

図１３ｂに示す出口側において、ＰＰＵは、トラフィックマネージャからパケットを受信した後（ステップ１３５０）、ローカルヘッダのＴＹＰＥフィールドをチェックする（ステップ１３５２）。パケットがＩＰパケット又はＣＰＵ宛のパケットであることをフィールドが示す場合、ＰＰＵは、パケットをＰＡＣＥに送信する（ステップ１３５４）。そうでない場合、以下に説明するように、ＰＰＵは、パケットに別の処理、例えば、仮想化及び変換処理を施す（ステップ１３５

６）。

【００８５】

前述のように、様々な状況において、ＳＰＵからＣＰＵにパケットが送られることになる。この状況は以下を含む。

１．宛先として記憶装置スイッチを有する非ＴＣＰパケット。このようなパケットは、本技術分野では理解されているように、ＩＣＭＰ、ＩＰ、ＲＩＰ、ＢＧＰ、又はＡＲＰパケットとすることができる。ＣＰＵは、スイッチ間通信及びＩＰ経路指定機能を実行する。また、パケットは、ＳＣＣに送られることになるＳＬＰ又はｉＳＮＳ要求であってもよい。

２．適切な経路指定宛先に一致するＣＡＭをもたないＩＰパケット。この状況は頻繁には発生しないはずであるが、発生した場合、ＣＰＵは、ＩＣＭＰパケットをソースＩＰアドレスに戻す。

３．非ｉＳＣＳＩ　ＴＣＰパケット。このパケットは、一般に、ＣＰＵがｉＳＣＳＩに関するＴＣＰセッションを確立又は終了させるためのものであり、典型的に、ＳＹＮ、ＦＩＮ、又はＲＳＴフラグ集合を有するパケットである。

４．非ＦＣＰ　ＦＣフレーム。このフレームは、ネームサービスのためのＦＬＯＧＩ、ＰＬＯＧＩ、及び他のＦＣＰ要求である。ｉＳＣＳＩ　ＴＣＰセッションと同様、これらのフレームによって、ＣＰＵは、ＦＣ装置を認識して通信することができる。１つの実施形態において、ＣＰＵは、ＳＣＣと通信してサービスを完了する必要がある。

５．ＳＣＳＩコマンド、応答、又はデータではないｉＳＣＳＩ　ＰＤＵ。このパケットは、ピング（ping）、ログイン、ログアウト、又はタスク管理とすることができる。一般に、セッションが完全に確立される前に別のｉＳＣＳＩ通信が必要とされる。ＣＰＵは、ログインを完了するためにＳＣＣデータベースからの情報を必要とすることになる。

６．読み出し／書込み／検査ではないＳＣＳＩコマンドを有するｉＳＣＳＩコマンドＰＤＵ。これらのコマンドは、仮想ターゲット動作が実行された場合にＣＰＵによって処理されることになるｉＳＣＳＩ制御コマンドである。

７．読み出し／書込み／検査ではないＳＣＳＩコマンドをもつＦＣＰフレーム。

これらのコマンドは、仮想ターゲット動作が実行された場合にＣＰＵによって処理されることになるＦＣＰ制御コマンドである。

【００８６】

仮想化：

前述のようにパケットが分類された後に、ＰＰＵは、ワイヤ速度で仮想化を行ない、且つ１つの実施形態においてはデータバッファリングをせずに行なう。受信された各々のパケットに関して、ＰＰＵは、パケットの種類（例えば、コマンド、Ｒ２Ｔ／ＸＦＲ＿ＤＲＹ、書込みデータ、読み出しデータ、応答、タスク管理／停止）を判定し、次に、入口側アルゴリズム（パケットがスイッチに入る場合）又は出口側アルゴリズム（パケットがスイッチを出る場合）のいずれかを実行して、仮想ターゲットを物理的ターゲットに、又は物理的ターゲットを仮想ターゲットに変換する。従って、仮想化機能は、入口側ポートと出口側ポートとの間に分散される。別のワイヤ速度処理をできるようにするために、ＣＡＭと共に仮想記述子を使用して、要求位置をアクセス位置にマッピングするようになっている。更に、各々のパケットに関しては特別の配慮がなされてもよい。例えば、パケットの宛先である仮想ターゲットは、複数の非連続領域にわたって区切ってもよく、ミラーリングを行ってもよく、又はその両方を行ってもよい。（ミラーリングに関しては、本明細書の「記憶サービス」のセクションで説明される）。以下に各々のパケット種類に関する入口側プロセス及び出口側プロセスを説明する。しかしながら、一般的に、各々のパケットに関する入口側プロセスは、仮想ターゲットを確認し、パケットの宛先である出口側ポートを決定し、応答パケットを追跡できるように追跡タグを残す。一般的に、出口側プロセスは、追跡タグを保持し続け、且つブロックアドレスの調節を行い、仮想世界から物理的世界に変換する。

【００８７】

コマンドパケット－入口側：

仮想ターゲットへの又はそこからの伝送タスクを開始するために、ＳＣＳＩコマンドは、それぞれｉＳＣＳＩ　ＰＤＵ又はＦＣＰ　ＩＵのｉＳＣＳＩ又はＦＣイニシエータによって常に送信される。図１４及び図１４ａを参照すると、このパ

ケットが（分類後に）ＰＰＵで受信されると（ステップ１４０２）、次に、ｉＳＣＳＩイニシエータの場合にはＴＣＰ制御ブロックインデックス及び論理ユニット番号（ＬＵＮ）を使用して、或いはＦＣイニシエータの場合にはＳ＿ＩＤ及びＬＵＮを使用して、有効なＶＴＤ　ＩＤが存在するか否かを判定するためにＰＰＵ　ＣＡＭをチェックする（ステップ１４０４）。各々の場合のＬＵＮは、それぞれｉＳＣＳＩ　ＰＤＵ又はＦＣＰ　ＩＵで見つけることができる。有効なＶＴＤ　ＩＤが見つからない場合、応答パケットがイニシエータに返送される（ステップ１４０６）。有効なＶＴＤ　ＩＤが見つかった場合、有効でないパケットに関するチェックが行われる（ステップ１４０８）。このようなチェックとしては、仮想ターゲットに関する未処理コマンド番号が最大許容番号を超えたか否か、又はアクセス要求を受けたブロックが許容範囲内にあるか否かを判定するためのチェックを挙げることができる。無効パラメータが存在する場合、ｉＳＣＳＩ又はＦＣイニシエータに応答パケットが返送される（ステップ１４０６）。

【００８８】

チェックした全パラメータが有効な場合、図１４ａに示すように、タスクインデックスは入口側タスク制御ブロック（ＩＴＣＢ）と共に割り当てられる（ステップ１４１０）。タスクインデックスはＩＴＣＢを指示又は識別する。ＩＴＣＢは、（ＶＴＤから取得された）フローＩＤ、（ｉＳＣＳＩパケット自体からの）ＶＴＤ　ＩＤ、ＣｍｄＳＮ、並びにｉＳＣＳＩ　ＰＤＵに送信されたイニシエータタスクタグ、又はＦＣＰフレームヘッダ内のＯＸ＿ＩＤを記憶する。ＩＴＣＢは、ＰＰＵ　ＳＲＡＭに記憶される。勿論、任意の所定時間に多数の処理中のコマンドが存在してもよいので、ＰＰＵは、任意の特定時間にＩＴＣＢ番号を記憶することができる。各々のＩＴＣＢは、それぞれのタスクインデックスによって参照されることになる。

【００８９】

ＶＴＤは、特定の仮想ターゲットに対する未処理コマンドを追跡するので、新しいＩＴＣＢが確立されると、未処理コマンド番号を増分する必要がある（ステップ１４１２）。特定の実施形態において、ＶＴＤは、特定の仮想ターゲットの任意の１つに対して未処理であろうコマンドの最大番号を設定する。フローＩＤ、

ＶＴＤ　ＩＤ、及びタスクインデックスの全ては、ローカルヘッダにコピーされる（ステップ１４１４）。フローＩＤは、トラフィックマネージャに宛先ラインカード及びポートを教える。その後、タスクインデックスは、パケットの特定のタスクを識別するために出口側ポートから返送されることになる。最後に、パケットは、トラフィックマネージャ、次に経路指定ファブリックへ送信されるので、最終的には出口側ＰＰＵに到達する（ステップ１４１６）。

【００９０】

仮想ターゲットが複数の領域で構成される場合、各々の領域に対して１つのＶＴＤで識別される複数のフローＩＤが存在することになる。ＰＰＵは、パケットに関するブロックアドレスをチェックし、次に、正しいフローＩＤを選択する。例えば、仮想ターゲットが２つの１Ｇｂ領域を有し、コマンドのブロックアドレスが第２の領域にある場合、ＰＰＵは、第２の領域のフローＩＤを選択する。換言すると、フローＩＤは、宛先／出口ポートを決定する。読み出しコマンドが領域境界を越える場合には、コマンドが、第１の領域で開始ブロックアドレスを指定し、第２の領域で終了ブロックアドレスを指定することを意味し、適切なデータを第１の領域から読取った後、ＰＰＵは、残りのブロックを読取るために第２の領域にコマンドを繰り返す。領域境界を越える書込みコマンドに関しては、ＰＰＵは、両方の領域にコマンドをコピーすると共に書込みデータの順番を管理する。読み出しコマンドが領域境界を越える場合、２つの領域に対する２つの読み出しコマンドが存在することになる。第２の読み出しコマンドは、第１の読み出しコマンドの完了後にのみ送信され、データがイニシエータに連続的に確実に返送されるようにする。

図１４ａに関して、必ずしもローカルヘッダ内のフィールドの全てが図示されていないことに留意されたい。

【００９１】

コマンドパケット－出口側：

図１５及び図１５ａを参照すると、出口側ポートに指定されたコマンドＰＤＵ又はＩＵは、スイッチファブリックを通過後に、ＰＰＵに到達する（ステップ１５０２）。次に、ＰＰＵは、パケットの宛先である物理的装置の識別を試みる（ス

テップ１５０４）。この識別を行うために、ローカルヘッダからのＶＴＤ　ＩＤを使用して、ＰＴＤ　ＩＤ（物理的ターゲット記述子識別子）に関するＰＰＵＣＡＭを検索する。ＶＴＤ　ＩＤは、特定の出口側ＰＰＵに関連する特定のＰＴＤ　ＩＤに関係付けされ、インデックス付けされる。ＰＴＤは、ＶＴＤと同様に、ＰＰＵ　ＳＲＡＭに記憶され、ＶＴＤで見出せるのと同様の情報を含む。検索が失敗した場合、これはＣＰＵによって直接送信されたコマンドパケットであり、ＰＰＵによってどんな追加の処理も要求されないとみなされ、結果的に、ＰＰＵは、ローカルヘッダのフローＩＤに基づいてパケットを適切な出口側ポートに送る。検索が成功した場合、ＰＴＤ　ＩＤは、仮想ターゲットがマッピングされ、かつパケットを現在処理中の特定の出口側ラインカードと通信状態にある物理的ターゲット（領域を含む）を識別する。

【００９２】

次に、図１５ａに示すように、ＰＰＵは、タスクインデックスを出口側タスク制御ブロック（ＥＴＣＢ）に割り当てる（ステップ１５０６）。１つの実施形態において、出口側に使用されるタスクインデックスは、入口側に使用されるものと同じである。また、タスクインデックスは、ＥＴＣＢを識別する。更に、ＥＴＣＢは、コマンドに必要な任意の他の制御情報を記憶するが、その制御情報としてはｉＳＣＳＩ　ＰＤＵのＣｍｄＳＮ、又はＦＣＰ　ＩＵの交換シーケンスを挙げることができる。

【００９３】

次に、ＰＴＤコンテンツを用いて、ＰＰＵは、仮想ターゲットからのＳＣＳＩブロックアドレスを物理的装置のブロックアドレスに変換する（ステップ１５０８）。仮想ターゲットのブロックアドレスを領域の開始ブロックオフセットに加算することによって、この変換を行うことができる。例えば、アクセスしようとする仮想ターゲットブロックが１９９０であり、且つ対応する第１の領域の開始オフセットが３０００である場合、アクセスされる領域のブロックアドレスは４９９０である。次に、ＰＰＵは、適切なｉＳＣＳＩ　ＣｍｄＳＮ又はＦＣＰシーケンスＩＤを生成し（ステップ１５１０）、ｉＳＣＳＩ　ＰＤＵ又はＦＣＰフレームヘッダに入れる。また、ＰＰＵは、必要であればＦＣＰフレームヘッダを構

築し（特定の実施形態において、入口側ＰＰＵは、ＦＣＰヘッダから必要な情報を読み出した後にＦＣＰヘッダを取り除くが、他の実施形態では、ＦＣＰヘッダをそのままにしておき、必要なフィールドを単純に更新又は変更することができる）、又はｉＳＣＳＩターゲットに送信されたパケットについては、ＴＣＰ制御ブロックインデックスはＰＴＤからローカルヘッダにコピーされる（ステップ１５１２）。更に、ＰＰＵは、ｉＳＣＳＩ又はＦＣＰヘッダに必要とされる任意のフラグ又は他の変数を与える。次に、完成したｉＳＣＳＩ　ＰＤＵ又はＦＣＰフレームは、ＰＡＣＥに送信され（ステップ１５１４）、ＰＡＣＥは、ローカルヘッダを取り除き（ステップ１５１６）、適切なポートにパケットを送る（ステップ１５１８）。

【００９４】

複数領域の仮想ターゲットに関して、各々の領域は、異なる開始オフセットをもつ。従って、コマンドを２つの領域の間で分割する必要がある場合、ＰＰＵは、適切なアドレスを決める必要がある。例えば、仮想ターゲットが表１で定義した２つの領域をもつと想定する。

【００９５】

表１

領域	1	2
開始オフセット	３０００	５０００
ブロックサイズ	２０００	２５００

【００９６】

３０個のブロックに関するアドレス１９９０で始まる仮想ターゲットにアクセスすることが望まれる場合、第１の領域のＰＰＵは、コマンドを１０個のブロックに関するアドレス４９９０に送る（５１２０バイトのデータ、１つの実施形態において、ブロックは５１２バイト）。第２の領域のＰＰＵは、２０個のブロックに関するコマンドをアドレス５０００に送信する（１０，２４０バイトのデータ）。換言すると、第１の領域のＰＰＵは、アクセスされるアドレスを第１の領域の開始オフセットに加算し（３０００＋１９９０）、次に、そのアドレスを全サイズから減算して（２０００－１９９０）、アクセスできるブロック数を決め

る。第２の領域のＰＰＵは、開始オフセット（５０００）で始まり、残りのブロック（２０）を加算する（５０００から５０１９）ことになる。別の実施例として、仮想ブロック２０２０にアクセスすることが望まれる場合、第２の領域のＰＰＵは、第２の領域（５０００）のオフセットを加算する前に、第１の領域のサイズ（２０００）を減算し、結果的にアドレス５０２０を得ることになる。

【００９７】

Ｒ２Ｔ又はＸＦＲ＿ＲＤＹ－入口側：

図１６及び図１６ａを参照すると、前述のようにコマンドがターゲット記憶装置に送信された後、コマンドが書込みコマンドである場合、Ｒ２Ｔ　ＰＤＵ又はＸＦＲ＿ＲＤＹ　ＩＵが、書込みデータを受け取る準備ができている場合に記憶装置から受信される（ステップ１６０２）。ＰＰＵは、イニシエータタスクタグ又はＯＸ＿ＩＤを使用することにより、対応するＥＴＣＢを識別する（ステップ１６０４）。特定の実施形態において、パケットのイニシエータタスクタグ又はＯＸ＿ＩＤは、タスクインデックスと同じであり、ＥＴＣＢを識別する。無効なイニシエータタスクタグ又はＯＸ＿ＩＤのために、ＰＰＵが有効なＥＴＣＢを識別できなかった場合、パケットは破棄される。そうでない場合、ＥＴＣＢを識別した状態で、ＰＰＵは、ＥＴＣＢから入口側タスクインデックス（出口側タスクインデックスと異なる場合）及びＶＴＤ　ＩＤを検索する（ステップ１６０６）。また、ＰＰＵは、ＰＴＤからフローＩＤを検索し、フローＩＤは、ＥＴＣＢにおいてＰＴＤ　ＩＤによって識別される。フローＩＤは、トラフィックマネージャに対して元のイニシエータ（入口側）ポートのラインカードを指示する。フローＩＤ、ＶＴＤ　ＩＤ、及びタスクインデックスは、パケットのローカルヘッダにコピーされる（ステップ１６０８）。最後に、パケットは、トラフィックマネージャ及びスイッチファブリックに送信される（ステップ１６１０）。

【００９８】

Ｒ２Ｔ又はＸＦＲ＿ＲＤＹ－出口側：

Ｒ２Ｔ又はＸＦＲ＿ＲＤＹパケットがスイッチファブリックから出ていった後、イニシエータ（特定のタスクの元のコマンドを開始したデバイス）に返送される途中で、ＰＰＵによって受信される（ステップ１７０２）。タスクインデックス

は、ＰＰＵに対するＩＴＣＢを識別し（ステップ１７０４）、ＩＴＣＢから元のイニシエータタスクタグ及びＶＴＤ　ＩＤを取得することができる。Ｒ２Ｔ／ＸＦＲ＿ＲＤＹＤ期待データ伝送長、又はＢＵＲＳＴ＿ＬＥＮフィールドは、ＩＴＣＢに記憶される（ステップ１７０６）。ローカルヘッダは、ＦＣＰ　Ｄ＿ＩＤ又はＴＣＰ接続のためのＴＣＰ制御ブロックインデックスで更新される（ステップ１７０８）。ＩＴＣＢに記憶されている、元のパケットからの記憶されたＳ＿ＩＤは、Ｄ＿ＩＤになることに留意されたい。必要であれば、ＦＣＰフレームヘッダが構築されるか、又はそのフィールドが更新される（ステップ１７１０）。宛先ポート番号は、フローＩＤの代わりにローカルヘッダで指定され（ステップ１７１２）、イニシエータタスクタグと一緒にＳＣＳＩ　ＰＤＣに入れられるか、又は、ＦＣ接続の場合に、ＲＸ＿ＩＤ及びＯＸ＿ＩＤがＦＣＰフレームに入れられる。また、ＰＰＵは、ＰＤＵ又はＦＣＰヘッダに入れる必要がある、他の任意のフラグ又は変数を入れる。パケットはＰＡＣＥに送られ（ステップ１７１４）、ＰＡＣＥは、ローカルヘッダからの発信ポートを識別する。次に、ローカルヘッダは取り除かれ（ステップ１７１６）、送信に適したポートに送られる（ステップ１７１８）。

【００９９】

コマンドが２つ又はそれ以上の領域に分割される場合、例えば、コマンドが１つの領域で始まり別の領域で終わる場合、ＰＰＵは、第１の領域へのデータ伝送が完了するまで、第２の領域のＲ２Ｔ又はＸＦＲ＿ＲＤＹを保持する必要があり、これによってイニシエータからの連続的なデータ伝送を確実に行うことができる。更に、第２の領域のＲ２Ｔ又はＸＦＲ＿ＲＤＹのデータオフセットは、第１の領域に伝送されたデータ量を加算することによって変更されることが必要となるであろう。表１の実施例を参照すると、コマンドが３０個のブロックに関するブロック１９９０にアクセスする場合、第２の領域のＲ２Ｔ又はＸＦＲ＿ＲＤＹのデータオフセットは、１０個のブロックを追加する必要があり、その結果、１１番目のブロックが第２の領域に伝送されることになる最初のブロックになる。

【０１００】

書込みデータパケット－入口側：

イニシエータは、Ｒ２Ｔ又はＸＦＲ＿ＲＤＹパケットの受信後に書込みデータパケットを返送する。図１８及び図１８ａを参照すると、書込みデータｉＳＣＳＩＰＤＵ又はＦＣ　ＩＵがイニシエータから受信されると（ステップ１８０２）、パケットが属するＩＴＣＢを識別する必要がある（ステップ１８０４）。通常、ＩＴＣＢは、特定の実施形態においてタスクインデックスと同じである、ＲＸ＿ＩＤ又はターゲットタスクタグを使用して識別することができる。更に、ＳＰＵは、受信パケットが順序通りであることを識別する。しかしながら、特定の実施形態において、イニシエータは、要求していないデータ、即ち、Ｒ２Ｔ又はＸＦＲ＿ＲＤＹの受信前に送信されたデータを伝送することになる。この場合、ＰＰＵは、特定の仮想ターゲットの未処理タスクを検索することによってＩＴＣＢを見つける必要がある。しかし、ＩＴＣＢが見つからなかった場合、パケットは破棄される。ＩＴＣＢが見つかった場合、伝送されることになるデータ総量がＩＴＣＢにおいて更新される（ステップ１８０６）。フローＩＤ及びタスクインデックスは、パケットのローカルヘッダに追加される（ステップ１８０８）。次に、パケットは、トラフィックマネージャに送られ、最終的にはスイッチファブリックに送られる（ステップ１８１０）。

【０１０１】

コマンドが２つの領域の間に分割される場合、コマンドは第１の領域で始まり第２の領域で終わるので、ＰＰＵは、特定のデータが属する領域を割り出して、データパケットを正しい出口側ラインカードに送る必要がある。ＰＰＵは、領域に対して正しいフローＩＤを設定する。第１の領域上のデータ伝送の完了後、ＰＰＵは、第２の領域のＲ２Ｔ又はＸＦＲ＿ＲＤＹを受信したか否かをチェックする。連続的な伝送を確実に行うために、第１の領域上のデータ伝送が完了するまで、データは第２の領域に送信されないことになる。

【０１０２】

書込みデータパケット－出口側：

図１９及び図１９ａを参照すると、（トラフィックマネージャ経由で）スイッチファブリックから書込みデータパケットを受信すると（ステップ１９０２）、パケットのＥＴＣＢを識別する必要がある（ステップ１９０４）。一般的に、ＥＴ

ＣＢは、ローカルヘッダのタスクインデックスを使用して識別することができる。ＥＴＣＢが識別されると、ＰＰＵは、ＥＴＣＢ内の情報を使用して、ＰＤＵ又はＦＣＰフレームヘッダに関するデータオフセット等の任意の別のフラグ及び変数と一緒に、適切なｉＳＣＳＩ　ＤａｔａＳＮ又はＦＣＰシーケンスＩＤを生成する（ステップ１９０６）。ローカルヘッダは、ＰＴＤからのＴＣＰ制御ブロックインデックス又はＦＣＰ　Ｄ＿ＩＤで更新される（ステップ１９０８）。また、ポート番号は、ローカルヘッダに追加される。完成したｉＳＣＳＩ　ＰＤＵ又はＦＣＰフレームは、ＰＡＣＥに送信され（ステップ１９１０）、ＰＡＣＥは、ローカルヘッダを取り除き（ステップ１９１２）、パケットを適切なポートに送る（ステップ１９１４）。

【０１０３】

コマンドが２つの領域の間に分割される場合、第２の領域に対するパケットのデータオフセットが調整される必要がある。表１の実施例を使用すると、コマンドが３０個のブロックに関する１９９０で始まる仮想アドレスにアクセスする必要がある場合、第２の領域に対する書込みデータパケットのデータオフセットは、実際にはイニシエータからの１１番目のブロック１１が第２の領域の最初のブロックなので、１０個のブロック分を減算する必要がある。

【０１０４】

読み出しデータパケット－入口側：

図２０及び図２０ａを参照すると、読み出しコマンドの受信後に、ターゲット装置は、読み出しデータパケットに応答することになり、読み出しデータパケットは、（ＰＡＣＥでの分類後に）ＰＰＵにて受信されることになる（ステップ２００２）。次に、パケットのＥＴＣＢは、ＯＸ＿ＩＤ又はイニシエータタスクタグを使用して識別される（ステップ２００４）。更に、ＰＰＵは、シーケンス番号を使用するか又はデータオフセットが昇順であることを検査することによって、パケットが順番に受信されたか否かを検査する（ステップ２００６）。パケットが順序通りではない場合、読み出しコマンドはエラー終了する。しかしながら、パケットが適切な順番になっている場合、ＶＴＤ　ＩＤ、タスクインデックス、及びフローＩＤは、ＥＴＣＢ及びＶＴＤから検索され、ローカルヘッダにコピー

される（ステップ２００８）。パケットは、トラフィックマネージャに送信され、最終的にはスイッチファブリックに送信される（ステップ２０１０）。

【０１０５】

読み出しデータパケットが領域境界を越える場合、第２の領域からのパケットのデータオフセットが変更される必要がある。通常、このオフセットは、フローＩＤが第２の領域からのパケットを識別することになるので、以下に説明するように出口側で行われる。更に、データを連続的に確実に返送するために、第２の領域への読み出しコマンドは、第１の領域からの読み出しが完了するまでは送信されないことになる。

【０１０６】

読み出しデータパケット－出口側：

図２１及び図２１ａを参照すると、ＰＰＵが読み出しデータパケットをスイッチファブリックから受信すると（ステップ２１０２）、パケットのＩＴＣＢは、通常、ローカルヘッダのタスクインデックスを使用して識別される（ステップ２１０４）。ＩＴＣＢから、ＰＰＵは、イニシエータタスクタグ又はＯＸ＿ＩＤを検索する（ステップ２１０６）。ＩＴＣＢの保存データを使用して、ＰＰＵは、適切なｉＳＣＳＩ　ＤａｔａＳＮ又はＦＣＰシーケンスＩＤ、並びにＰＤＵ又はＦＣＰフレームヘッダの他のフラグ又は変数を生成する（ステップ２１０８）。ローカルヘッダは、ＶＴＤからのＴＣＰ制御ブロックインデックス又はＦＣＰ　Ｓ＿ＩＤで更新される（ステップ２１１０）。しかしながら、イニシエータに返送されるパケットに関して、元のパケットからのＳ＿ＩＤがＤ＿ＩＤとして使用されることに留意されたい。また、発信ポート番号は、ローカルヘッダに追加される。次に、パケットは、ＰＡＣＥに送信され（ステップ２１１２）、ＰＡＣＥは、ローカルヘッダを取り除き（ステップ２１１４）、パケットを適切なポートに送る（ステップ２１１６）。

【０１０７】

コマンドが２つの領域の間に分割される場合（ＩＴＣＢで追跡した事実）、第２の領域からのパケットのデータオフセットは、前述の方法と同様の方法で変更する必要がある。

【０１０８】

応答パケット－入口側：

図２２及び図２２ａを参照すると、応答パケットは、ターゲット装置から受信されることになる。次に、パケットに関するＥＴＣＢは、パケットのイニシエータタスクタグ又はＯＸ＿ＩＤを使用して識別される（ステップ２２０４）。特定の実施形態において、イニシエータタスクタグ又はＯＸ＿ＩＤは、タスクインデックスと同じものとなる。ＥＴＣＢが見つからなかった場合、パケットは破棄される。しかしながら、ＥＴＣＢが見つかった場合、タスクインデックスは、ＶＴＤＩＤ及びフローＩＤと一緒にパケットローカルヘッダにコピーされる（ステップ２２０６）。パケットは、トラフィックマネージャに送信され、最終的にはスイッチファブリックに送信される（ステップ２２０８）。最後に、応答パケットはタスクの完了を知らせるので、タスクのＥＴＣＢは解放される（ステップ２２１０）。

【０１０９】

応答パケット－出口側：

図２３及び図２３ａを参照すると、応答パケットは、スイッチファブリックを通過した後に、出口側ＰＰＵによって受信されることになる（ステップ２３０２）。パケットに関するＩＴＣＢは、ローカルヘッダからのタスクインデックスを使用して識別される（ステップ２３０４）。ＩＴＣＢが見つからなかった場合、パケットは破棄される。ＩＴＣＢが見つかった場合、仮想ターゲットに関する未処理コマンドカウントがＶＴＤにおいて減分される（ステップ２３０６）。ＰＰＵは、ＬＵＮ、ｉＳＣＳＩ　ＥｘｐＳｔａｔＳＮ、又はＦＣＰシーケンスＩＤをＩＴＣＢ内の情報から生成し、必要であれば、適切なＦＣＰヘッダを構築又は更新する（ステップ２３０８）。また、ＰＰＵは、ＰＤＵ又はＦＣフレームヘッダのために他のフラグ及び変数を構築する。ＰＰＵは、ＶＴＤから検索されたＴＣＰ制御ブロックインデックス又はＦＣＰ　Ｓ＿ＩＤ（これはＤ＿ＩＤになる）を用いてローカルヘッダを更新する（ステップ２３１０）。パケットは、ＰＡＣＥに送られ（ステップ２３１２）、ＰＡＣＥは、ローカルヘッダを取り除き（ステップ２３１４）、パケットを適切なポートに送る（ステップ２３１６）。

ＰＰＵは、ＩＴＣＢを解放する（ステップ２３１８）。

【０１１０】

書込みコマンドが２以上の領域に送信されている場合、応答パケットは、全領域に対する書込みが完了するまでイニシエータへ送信されない。

【０１１１】

図９から図２３の全てについて、種々のステップが特定の順番で実行されるように説明されているが、他の実施形態において、特定のステップの順番を変更してもよく、特定のステップを同時に実行してもよいことに留意されたい。

【０１１２】

タスク管理ＰＤＵ、異常終了、異常終了シーケンス／交換－入口側：

ＡＢＯＲＴ　ｉＳＣＳＩ機能、又は異常終了シーケンス／交換は、コマンドを通常とは異なる方法で終了する。ＰＰＵは、パケットのＯＸ＿ＩＤ又はイニシエータタスクタグを使用してＩＴＣＢを見つける。ＩＴＣＢが見つからなかった場合、コマンドは、既に完了したか又は受信されなかったと想定され、ＴＡＳＫ－ＮＯＴ－ＦＯＵＮＤを示す応答が生成されることになる。ＡＢＯＲＴがターゲット装置から受信された場合、ＰＰＵは、ＥＴＣＢを見つけて解放する。ＡＣＫはターゲット装置に返送され、ＡＢＯＲＴは、コマンドを終了するためにイニシエータに接続しているラインカードに送られる。ＡＢＯＲＴがイニシエータから受信された場合、ＡＢＯＲＴは、コマンドを終了するためにターゲットに接続しているラインカードに送られる。ＰＰＵは、それぞれのタスク制御ブロック、ＩＴＣＢ及びＥＴＣＢを解放する。

【０１１３】

タスク管理ＰＤＵ、異常終了、異常終了シーケンス／交換－出口側：

入口側ラインカードからのＡＢＯＲＴは、出口側ラインカードに対して、ＡＢＯＲＴをターゲット装置に送信することを指示する。完了応答がターゲットから返送される場合、ＥＴＣＢは解放される。ＥＴＣＢが見つからなかった場合、ＡＢＯＲＴは無視される。

【０１１４】

変換：

前述のように、本発明による記憶装置スイッチは、複数のプロトコルのいずれかに基づいてデータを送信する装置に結合することができる。また、前述のように、１つの実施形態において、サーバ及び記憶装置が利用するプロトコルは、ｉＳＣＳＩ及びファイバ・チャンネルである。しかしながら、スイッチが第１のプロトコルに基づいて作動するサーバに結合され、第２のプロトコルに基づいて作動する記憶装置に結合される場合、又は逆の場合も同様に、スイッチは、プロトコル変換を行う必要がある。従来、このような変換を行うためには、従来のシステムが仮にプロトコル変換を行うことができたとしても、パケットをメモリに記憶して転送前にＣＰＵによって操作する必要がある。対照的に、本発明による記憶装置スイッチは、スイッチにおいてパケットのバッファリングを全く行うことなくプロトコル変換を行うことができる。

【０１１５】

ｉＳＣＳＩ　ＰＤＵ及びファイバ・チャンネル　ＩＵの両方は、それぞれのパケット又はフレームにＳＣＳＩ　ＣＤＢ（コマンド記述子ブロック）を伝送するように設計される。従って、これらのプロトコルは、本発明の発明者が認識しているのと同様の意味論をもつ。表２は、各プロトコル間の比較を示す。

【０１１６】

表２

ＳＣＳＩフェーズ	ｉＳＣＳＩプロトコル	ＦＣプロトコル
仲裁（Arbitrate）及び選択	イーサネットパケットを送る。	ファイバ・チャンネルフレームを送信する
コマンド	コマンドＰＤＵ	コマンドフレーム
切断（Disconnect）	パケットを受信する	フレームを受信する
データ伝送のための再接続	Ｒ２Ｔ　ＰＤＵ	ＸＦＲ＿ＲＤＹフレーム
データ	ＴＣＰセグメント内のデータＰＤＵ	フレーム内のデータシーケンス
状態	応答ＰＤＵ	応答フレーム
異常終了及びリセット	ｉＳＣＳＩタスク管理	ファイバ・チャンネル　ＥＬＳ
待ち行列満杯状態	ＭａｘＣｍｄＳＮウィンドウ	タスク設定満杯
セッションログインなし	ｉＳＣＳＩログイン及びログアウト	ＰＬＯＧＩ及びＬＯＧＯ

【０１１７】

表２から、ｉＳＣＳＩコマンドＰＤＵとＦＣコマンドフレーム、Ｒ２Ｔ　ＰＤＵ

とＸＦＲ＿ＲＤＹフレーム、データＰＤＵとデータフレーム、応答ＰＤＵと応答フレームとの間には相関関係があることが分かる。このような相関関係は、ＰＰＵにおいて、以下に説明するように、バッファリングを行うことなく１つのパケットから別のパケットへフィールドをマッピングすることで行われる直接的な変換に適する。異常終了及びリセット、セッションログイン及びログアウト、及び待ち行列満杯状態は、他のパケットに関連して不定期に起こり、ラインカードのＣＰＵに送られて処理される（ＰＰＵによって実行される、ＳＣＳＩデータ移動（読み出し／書込み）コマンドの異常終了を除く）。ＳＣＳＩの仲裁及び選択、並びに切断に関して、ｉＳＣＳＩ及びＦＣの両者は、単純にパケット／フレームを送受信することに留意されたい。

【０１１８】

パケットがＰＰＵへ到着すると、仮想化と同様に、ＰＰＵは、ＣＡＭを検索して受信コマンドが特定のセッション（ｉＳＣＳＩ又はＦＣのいずれか）及び特定の仮想ターゲットに属するか否かを判定し、パケットに関連するＶＴＤを識別する。ＣＡＭ検索は、前述のように、ＴＣＰ制御ブロックインデックス及びＬＵＮ（ｉＳＣＳＩパケットの場合）、又はＳ＿ＩＤ及びＬＵＮ（ＦＣフレームの場合）を使用して行われる。しかしながら、本発明の１つの実施形態において、変換は、出口側ＰＰＵ（スイッチファブリックを通過した後のパケットを受信するＰＰＵ）にて行われる。また、出口側ＰＰＵはＣＡＭを検索するが、パケットのローカルヘッダのＶＴＤ　ＩＤを使用してＰＴＤを見つける。

【０１１９】

仮想化及び変換機能の両方に関して説明されているが、各種機能に関して説明した他のステップと同様に、ＣＡＭ検索は、ＰＰＵによって１度だけ行われる必要があり、また、説明した全機能（例えば、分類、仮想化、及び変換）に対して行われる各種ステップは、多くの点で統合できる点に留意されたい。

【０１２０】

同様に、仮想化機能に関して前述したように、ＶＴＤは、仮想ターゲット及び物理的ターゲットに関する変数を追跡し続けるが、同様に、ＰＰＵは、典型的にはＩＴＣＢ及びＥＴＣＢ（ＳＣＳＩコマンド毎に各々１つ）の各プロトコル間で共

有されない変数を追跡し続ける。このような変数としては、ｉＳＣＳＩに関してはタスクタグ、ＣｍｄＳＮ、ＤａｔａＳＮ、及びＳｔａｔＳＮであり、ファイバ・チャンネルに関してはＯＸ＿ＩＤ、ＲＸ＿ＩＤ、交換シーケンス番号、及びシーケンス開始フラグを挙げることができる。ＰＰＵが、ＶＴＤ（又はＰＴＤ）並びにそれぞれのＥＴＣＢ又はＩＴＣＢを有すると、変換を行うのに必要な全ての情報を有する。ｉＳＣＳＩからＦＣへの変換、又はその逆の変換は、一般的に、受信パケット（例えば、ｉＳＣＳＩ）のフィールドから情報を取得すること、及び情報を送信パケット（例えば、ＦＣＰ）の対応フィールドへマッピングすることを必要とする。

【０１２１】

ＦＣターゲットに対するｉＳＣＳＩイニシエータ：

まず、ｉＳＣＳＩイニシエータ（サーバ）からＦＣターゲット（記憶装置）への変換について説明する。ｉＳＣＳＩコマンドＰＤＵからＦＣＰ＿ＣＭＮＤ　ＩＵへの変換は、以下の表３に基づいて行われる。また、図８ａ～図８ｉを参照されたい。

【０１２２】

表３

ｉＳＣＳＩコマンドＰＤＵから	ＦＣＰ＿ＣＭＮＤ　ＩＵへ
ｉＳＣＳＩ　ＰＤＵのＬＵＮフィールド	ＦＣＰ＿ＬＵＮ
ＡＴＴＲ（３ビット）	ＦＣＰ＿ＣＮＴＬ
ＣＤＢフィールド	ＦＣＰ＿ＣＤＢ
期待データ伝送長	ＦＣＰ＿ＤＬ
	ＯＸ＿ＩＤ、ＳＥＱ＿ＩＤ、ＳＥＱ＿ＣＮＴ

【０１２３】

表３によれば、ｉＳＣＳＩ　ＰＤＵのＬＵＮフィールドの内容は、ＦＣＰ＿ＣＭＮＤ　ＩＵのＦＣＰ＿ＬＵＮフィールドにマッピングされる。物理的ターゲットのＬＵＮは、ＰＴＤから取得される。ｉＳＣＳＩタスク属性フィールドＡＴＴＲの３ビットのみが、ＦＣＰ＿ＣＮＴＬフィールドにマッピングされる。ｉＳＣＳＩ　ＰＤＵのＣＤＢフィールドの内容は、ＦＣＰ＿ＣＤＢフィールドにマッピングされる。データ伝送サイズフィールドの内容は、ＦＣＰ＿ＤＬフィールドにマ

ッピングされる。ＯＸ＿ＩＤは、ＦＣＰフレームヘッダに固有なので、ターゲットからの各々のパケットの識別を容易にするために、典型的にＥＴＣＢからのタスクインデックスを用いて、ＰＰＵによって情報が与えられる。ＦＣＰフレームヘッダの他のフィールドは、ＰＴＤ又はＶＴＤからの情報を用いて簡単に生成することができる。

【０１２４】

ＦＣ記憶装置が応答する場合、ＦＣ記憶装置は、ＦＣ　ＸＦＲ＿ＲＤＹフレームで応答することになり、ＦＣ　ＸＦＲ＿ＲＤＹフレームは、ｉＳＣＳＩ　Ｒ２Ｔ　ＰＤＵに再変換する必要がある。

【０１２５】

表４

ＦＣＰ　ＸＦＲ＿ＤＲＹから	Ｒ２Ｔ　ｉＳＣＳＩ　ＰＤＵへ
ＤＡＴＡ＿ＲＯ	バッファオフセット
ＢＵＲＳＴ＿ＬＥＮ	データ伝送長
	イニシエータタスクタグ及び他のフィールド

【０１２６】

表４に示すように、バッファオフセット及びデータ伝送長フィールドは、ＦＣ　ＸＦＲ＿ＲＤＹフレームから直接マッピングされることができる。しかしながら、ＳｔａｔＳＮ、ＥｘｐＣｍｐＳＮ、ＭａｘＣｍｄＳＮ、及びＲ２ＴＳＮ等の他のフィールドをＩＴＣＢから取得する必要がある。更に、ｉＳＣＳＩ　Ｒ２Ｔ　ＰＤＵ固有のタスクタグのような変数は、通常、ＰＴＤ又はＶＴＤからのフィールドを使用して、ＰＰＵによってパケットに入れられる。

【０１２７】

Ｒ２Ｔの受信後、ｉＳＣＳＩイニシエータは、書込みデータＰＤＵを送信することになり、書込みデータＰＤＵは、ＦＣＰ　Ｄａｔａ　ＩＵに変換する必要がある。

【０１２８】

表５

ｉＳＣＳＩ読み出しデータＰＤＵから	ＦＣＰ　ＤＡＴＡ　ＩＵへ
バッファオフセット	ＲＬＶＴ＿ＯＦＦ
ペイロード	ペイロード
	ＯＸ＿ＩＤ、ＳＥＱ＿ＣＮＴ

【０１２９】

表５に示すように、ＦＣＰ　ＤＡＴＡ　ＩＵのＲＬＶＴ＿ＯＦＦフィールドは、ｉＳＣＳＩ　ＰＤＵのバッファオフセットフィールドからマッピングされることになる。各々のパケット／フレームに関するペイロードは、全く同じである。更に、ＥＴＣＢから取得した、ＯＸ＿ＩＤ及びＳＥＱ＿ＣＮＴ等のＦＣフレーム固有の変数が追加される。

【０１３０】

ｉＳＣＳＩイニシエータから最初に送信されたｉＳＣＳＩコマンドが読み出しデータコマンドである場合、ＦＣターゲットは、ＦＣＰ＿ＤＡＴＡ　ＩＵで応答することになり、ＦＣＰ＿ＤＡＴＡ　ＩＵは、ｉＳＣＳＩ読み出しデータＰＤＵに変換する必要がある。

表６

ＦＣＰ　ＤＡＴＡ　ＩＵから	ｉＳＣＳＩ読み出しデータＰＤＵへ
ＲＬＶＴ＿ＯＦＦ	バッファオフセット
データペイロード	データペイロード
	イニシエータタスクタグ、残りカウント

【０１３１】

表６に示すように、ｉＳＣＳＩ　ＰＤＵのバッファオフセットは、ＦＣＰ　ＩＵのＲＬＶＴ＿ＯＦＦフィールドからマッピングされることになる。他の全フィールドは、ＩＴＣＢ並びにタスクタグ等のＰＤＵ固有の変数から取得される。

【０１３２】

タスクが完了すると（例えば、データの読み出し又は書込みが終了すると）、ＦＣＰターゲットは、ｉＳＣＳＩフォーマットに変換する必要がある応答パケット（ＦＣＰ＿ＲＳＰ　ＩＵ）を送信する。

表７

ＦＣＰ応答ＩＵから	ｉＳＣＳＩ応答ＰＤＵへ
ＦＣＰ＿ＳＴＡＴＵＳ	フラグ及び状態フィールド
ＦＣＰ＿ＳＮＳ＿ＬＥＮ	データセグメント長
ＦＣＰ＿ＲＥＳＩＤ	基本残余カウント
ＦＣＰ＿ＳＮＳ＿ＩＮＦＯ	センスデータ
ＦＣＰ＿ＲＳＰ＿ＩＮＦＯ	エラーコード
	イニシエータタスクタグ、ＭａｘＣｍｄＳＮ、ＥｘｐＣｍｄＳＮ

【０１３３】

表７に示すように、ＦＣ　ＩＵの状態フィールドは、ｉＳＣＳＩ　ＰＤＵのフラグ及び状態フィールドにマッピングされる。ＦＣＰ＿ＳＮＳ＿ＬＥＮ、ＦＣＰ＿ＲＥＳＩＤ、及びＦＣＰ＿ＳＮＳ＿ＩＮＦＯは、それぞれ、データセグメント長、基本残余カウント、及びセンスデータにマッピングされる。ＦＣＰ＿ＲＳＰ＿ＩＮＦＯフィールドは、ｉＳＣＳＩエラーコードにマッピングされる必要がある伝送エラー用である。最後に、ｉＳＣＳＩ状態ＰＤＵに固有のタスクタグ、又はＥｘｐＣｍｄＳＮ、ＳｔａｔＳＮ、ＭａｘＣｍｄＳＮ、ＥｘｐＤａｔａＳＮ、ＥｘｐＲ２ＴＳＮ等の変数はＩＴＣＢ又はＶＴＤから追加される。

【０１３４】

異常終了タスクセット等のタスク管理用のＦＣＰ＿ＣＮＴＬにフラグがある場合、別個のｉＳＣＳＩタスク管理コマンドは、ｉＳＣＳＩイニシエータ装置に送信されることになる。同様に、ｉＳＣＳＩタスク管理ＰＤＵが受信された場合、ＦＣＰ＿ＣＮＴＬに適切なフラグを有するＮＯＰ　ＰＦコマンドがターゲット装置に送信される。

【０１３５】

前記の表には、ｉＳＣＳＩ　ＰＤＵ又はＦＣＰフレームのいずれかに固有の全フィールドが記載されていない点に留意されたい。フィールドを完全に記載するために図８ａから図８ｉを参照することができる。記載されていない任意のフィールドについては、関連のタスク制御ブロック、ＶＴＤ、ＰＴＤから取得することができ、又は簡単に生成することができることを理解されたい（例えば、ＦＣＰ形式フィールドは、常に０ｘ０８である）。

【０１３６】

ｉＳＣＳＩターゲットに対するＦＣイニシエータ：

ＦＣＰからｉＳＣＳＩへの変換は、ｉＳＣＳＩからＦＣＰへの変換の逆である。この場合も変換は出口側ＰＰＵで行われる。最初に、ＦＣＰイニシエータはＦＣＰコマンドを送信することになるが、ＦＣＰコマンドは、ｉＳＣＳＩターゲットに適するように変換する必要がある。

【０１３７】

表８

ＦＣＰコマンドＩＵから	ｉＳＣＳＩコマンドＰＤＵへ
ＦＣＰ＿ＬＵＮ	ＬＵＮ
ＦＣＰ＿ＣＮＴＬ	ＡＴＴＲ
ＦＣＰ＿ＣＤＢ	ＣＤＢ
ＦＣＰ＿ＤＬ	期待データ伝送長
	ＣｍｄＳＮ、タスクタグ、ＥｘｐＳｔａｔＳＮ

【０１３８】

表８に示すように、ＦＣ　ＩＵのＬＵＮ、ＣＮＴＬ、ＣＤＢ、及びＤＬフィールドは、ｉＳＣＳＩ　ＰＤＵのＬＵＮ、ＡＴＴＲ、ＣＤＢ、及びデータ伝送サイズフィールドにマッピングされる。更に、ＣｍｄＳＮ及びタスクタグ等のｉＳＣＳＩ　ＰＤＵ固有の変数は、ＰＰＵによって作成され、ＣｍｄＳＮ及びタスクタグの両方は、ＥＴＣＢから取得することができる。データセグメント長フィールドは、ＦＣＰフレームに関する当面のデータがないためにゼロになる。

【０１３９】

ｉＳＣＳＩターゲットがコマンドを受信した後（コマンドは書込みコマンド）、ターゲットは、Ｒ２Ｔ　ＰＤＵで応答することになり、Ｒ２Ｔ　ＰＤＵは、ＦＣＰ　ＸＦＲ＿ＲＤＹ　ＩＵに変換する必要がある。

【０１４０】

表９

ｉＳＣＳＩ　Ｒ２Ｔ　ＰＤＵから	ＦＣＰ　ＸＦＲ＿ＲＤＹ　ＩＵへ
バッファオフセット	ＤＡＴＡ＿ＲＯ
データ伝送長	ＢＵＲＳＴ＿ＬＥＮ
	ＲＸ＿ＩＤ、ＳＥＱ＿ＩＤ

【０１４１】

表９に示すように、ｉＳＣＳＩ　ＰＤＵのバッファオフセット及びデータ伝送長フィールドは、ＸＦＲ＿ＲＤＹ　ＩＵのＤＡＴＡ＿ＲＯ及びＢＵＲＳＴ＿ＬＥＮフィールドにマッピングされる。更に、ＰＰＵは、ＩＴＣＢで利用可能なＲＸ＿ＩＤ及びＳＥＱ＿ＩＤ等のＦＣＰ　ＩＵに固有の変数を追加する。

【０１４２】

ＦＣイニシエータは、ＸＦＲ＿ＲＤＹ　ＩＵの受信後に、ｉＳＣＳＩフォーマットに変換する必要がある書込みデータを送信することになる。

【０１４３】

表１０

ＦＣＰデータＩＵから	ｉＳＣＳＩ書込みデータＰＤＵへ
ＲＬＶＴ＿ＯＦＦ	バッファオフセット
ペイロード	ペイロード
	ＤａｔａＳＮ、ＥｘｐＣｍｄＳＮ、ターゲットタスクタグ

【０１４４】

表１０に示すように、書込みデータに関して、ＦＣＰ　ＩＵのＲＬＶＴ＿ＯＦＦは、ｉＳＣＳＩ　ＰＤＵのバッファオフセットにマッピングされるが、各々のペイロードは同じである。更に、他のフィールドは、ｉＳＣＳＩデータＰＤＵに固有のＤａｔａＳＮ等の変数も含めてＥＴＣＢから取得される。

【０１４５】

元のイニシエータコマンドが読み出しコマンドである場合、ｉＳＣＳＩターゲットは、ＦＣＰフォーマットで入れる必要がある読み出しデータを用いて応答することになる。

【０１４６】

表１１

ｉＳＣＳＩ読み出しデータＰＤＵから	ＦＣＰ　ＤＡＴＡ　ＩＵへ
バッファオフセット	ＲＬＶＴ＿ＯＦＦ
ペイロード	ペイロード
	ＲＸ＿ＩＤ、ＳＥＱ＿ＩＤ

【０１４７】

表１１に示すように、バッファオフセットフィールドは、ＦＣＰ　ＩＵのＲＬＶＴ＿ＯＦＦフィールドにマッピングされるが、両者のペイロードは同じである。更に、ＰＰＵは、ＩＴＣＢに見つけることができる、ＲＸ＿ＩＤ及びＳＥＱ＿ＩＤ等のＦＣＰ　ＩＵに固有の変数を追加する必要がある。

【０１４８】

最後に、タスクが完了すると、ｉＳＣＳＩターゲットは、応答ＰＤＵを送信することになり、応答ＰＤＵは、ＦＣＰ　ＲＳＰ　ＩＵに変換する必要がある。

【０１４９】

表１２

ｉＳＣＳＩ応答ＰＤＵから	ＦＣＰ　ＲＳＰ　ＩＵへ
フラグ及び状態	ＦＣＰ＿ＳＴＡＴＵＳ
データセグメント長	ＦＣＰ＿ＳＮＳ＿ＬＥＮ
基本残余カウント	ＦＣＰ＿ＲＥＳＩＤ
センスデータ	ＦＣＰ＿ＳＮＳ＿ＩＮＦＯ
伝送エラー	ＦＣＰ＿ＲＳＰ＿ＩＮＦＯ
	ＯＸ＿ＩＤ、ＳＥＱ＿ＩＤ

【０１５０】

表１２に示すように、ｉＳＣＳＩ　ＰＤＵのフラグ及び状態は、ＦＣＰ　ＩＵのＳＴＡＴＵＳフィールドにマッピングされる。ｉＳＣＳＩフィールドのデータセグメント長、基本残余カウント、及びセンスデータの全ては、ＦＣＰ　ＩＵのＦＣＰ＿ＳＮＳ＿ＬＥＮ、ＦＣＰ＿ＲＥＳＩＤ、及びＦＣＰ＿ＲＳＰ＿ＩＮＦＯフィールドにそれぞれマッピングされる。伝送エラーは、ＦＣＰ　ＩＵのＦＣＰ＿ＲＳＰ＿ＩＮＦＯフィールドにマッピングされる。更に、ＯＸ＿ＩＤ及びＳＥＱ＿ＩＤ等のＦＣＰ　ＩＵに固有の変数は、ＰＰＵによって追加される。

【０１５１】

異常終了タスクセット等のｉＳＣＳＩタスク管理パケットを受信した場合、これはＦＣＰ＿ＣＮＴＬフィールドにタスク管理フラグをもつＮＯＰ　コマンドを使用してＦＣ装置に送信されることになる。

【０１５２】

前記の表には、ｉＳＣＳＩ　ＰＤＵ又はＦＣＰフレームのいずれかに固有の全フ

ィールドが記載されていない点に留意されたい。フィールドを完全に記載するために図８ａ～図８ｉを参照することができる。記載されていない任意のフィールドについては、関連のタスク制御ブロック、ＶＴＤ、ＰＴＤから取得することができ、又は簡単に生成することができることを理解されたい（例えば、ＦＣＰ形式フィールドは、常に０ｘ０８である）。

【０１５３】

記憶サービス：

本発明の実施形態によるスイッチは、タスクを複数のラインカード上に分散することによって、スイッチ型の記憶サービスをワイヤ速度で行うので、スループットを最大にすることができる。本発明の１つの実施形態で提供される記憶サービスとしては、ローカルミラーリング、低速リンク上でのミラーリング、スナップショット、仮想ターゲットクローニング（複製）、第三者コピー、定期的スナップショット及びバックアップ、及びリストア等を挙げることができる。これらのサービスの各々について以下で更に詳細に説明する。他の実施形態では、これよりも多い又は少ないサービスを行うことができる。

【０１５４】

特定のサービスを説明する前に、図２４を参照すると、概括的に、記憶サービスは、最初に記憶装置スイッチとのイーサネット接続上の管理ステーション（又は他の装置）によって起動される（ステップ２４０２）。このようなイーサネット通信は、１つの実施形態においてＳＣＣ６１０（図６）で行われる。ＳＣＣは、データベースを通じて、そのサービスのためのラインカードを判定し、そしてＶＴＤ及びＬＵＮ情報を含む全ての関連情報をこれらのラインカードに送り、そのサービスを実行する（ステップ２４０４）。ＳＣＣがラインカード毎に有する全ての情報は、イーサネット通信上でのカード間通信を使用して、ＳＣＣからラインカードへ送られる。次に、ラインカードは、要求された実際のサービスを行う（ステップ２４０６）。タスクが完了すると、ＳＣＣは、管理ステーションへの返送応答を開始して（ステップ２４０８）、サービスの完了を指示する。従って、従来のシステムとは異なり、管理ステーションは、サービス要求を開始する以外は、そのサービスに関与する必要は全くない。

【０１５５】

ローカルミラーリング：

仮想ターゲットがミラーリングされた場合、即ち、そのデータと全く同じコピーが２つの別個の物理的位置に記憶された場合、ミラーリングされた仮想ターゲットの「メンバ」と呼ばれる場合が多い。ＶＴＤ内のフローＩＤは、パケットが複数の出口側ポートにマルチキャストされることを表す。ミラーリングされた仮想ターゲットにおいて、書込みコマンドが領域境界を越える場合、ＰＰＵは、ミラーリングされたターゲットの各々のメンバに対する各々の領域に関するパケットを複製することになる。また、ＰＰＵは、適切なフローＩＤをトラフィックマネージャに送り、トラフィックマネージャは、受信した各々のコマンドを複数の出口側ポートへ送信する。ミラーリングされた仮想ターゲットからの読み出し時に、ＰＰＵは、最小平均応答時間を有するミラーリングされたターゲットの１つのメンバを選択する。そのメンバのフローＩＤは、読み出しコマンドを選択された出口側ポートに導く。応答時間は、ＶＴＤで利用可能である。

【０１５６】

書込みコマンドの送信後に、ミラーリングされたターゲットのメンバの１つからＲ２Ｔ又はＸＦＲ＿ＲＤＹを受信した場合、ＰＰＵは、全てのメンバ及び／又は領域がＲ２Ｔ又はＸＦＲ＿ＲＤＹを返送するまで待機する。全てのメンバが応答済みになると、ＰＰＵは、データを受信するために利用可能な最小ブロックを指定するＲ２Ｔ又はＸＦＲ＿ＲＤＹを、イニシエータに送信する準備を行うことになる。即ち、データが返送されると、データは全てのミラーリングされたメンバにマルチキャストされることになるが、メンバは、要求した以上のデータを受信することはできない。従って、ＰＰＵは、ＩＴＣＢにおいて、各々の領域に関するＲ２Ｔ又はＸＦＲ＿ＲＤＹで指定された要求データ量を追跡する必要もある。最小量のデータが受信され（イニシエータから）、ミラーリングされたターゲットの各々のメンバにマルチキャストされると、ＰＰＵは、最小量のデータを要求した領域が別のＲ２Ｔ又はＸＦＲ＿ＲＤＹを送信するのを待つ。２つの（又はそれ以上の）ターゲットが最小量のデータを要求した場合（即ち、両者が同じデータ量を要求した場合）、ＰＰＵは、最小量を要求した両方の（又は全ての）ター

ゲットが別のＲ２Ｔ又はＸＦＲ＿ＲＤＹを送信するまで待機する。次に、ＰＰＵは、全ての領域の最小残量のＲ２Ｔ又はＸＦＲ＿ＲＤＹを返送する。このプロセスは、全領域が全ての要求データをもつまで続く。１つの実施例を表１３に示す。

【０１５７】

表１３

	領域１	領域２	イニシエータへ
書込まれる全データ	４ｋ	４ｋ	
第１のＲ２Ｔ又はＸＦＲ＿ＲＤＹで指定されたサイズ	２ｋ	３ｋ	
イニシエータからのＰＰＵ要求			２ｋ
不満足なＲ２Ｔ又はＸＦＲ＿ＲＤＹ（２ｋ書込み後）	０ｋ	１ｋ	
第２のＲ２Ｔ又はＸＦＲ＿ＲＤＹで指定されたサイズ	２ｋ		
イニシエータからのＰＰＵ要求			１ｋ
不満足なＲ２Ｔ又はＸＦＲ＿ＲＤＹ（１ｋ書込み後）	１ｋ	０ｋ	
第３のＲ２Ｔ又はＸＦＲ＿ＲＤＹで指定されたサイズ		１ｋ	
イニシエータからのＰＰＵ要求			１ｋ
不満足なＲ２Ｔ又はＸＦＲ＿ＲＤＹ（１ｋ書込み後）	０ｋ	０ｋ	

【０１５８】

低速リンク上のリモートミラーリング：

前述のように、ミラーリングは、２つの同じデータ集合の各々が、別個の物理的位置にそれぞれ記憶される場合に起こる。大部分の従来システムは、ローカルミラーリング、即ち、同一ＳＡＮ上に存在する各装置におけるミラーリングだけをサポートする。しかしながら、本発明の実施形態は、低速リンク上でのミラーリングをサポートする。例えば、データの１つのコピーが１つのＳＡＮ上にあり、データの第２のコピーがＳＡＮから離れた位置に、例えば、第２のＳＡＮ上に記憶されている場合のミラーリングをサポートする。例えば、図４を参照すると、データのローカルコピーがＳＡＮ４０２にあるが、リモートミラーコピーは、ＳＡＮ４０４にあることができる。従って、リモートミラーリングは、本発明の実施形態のスイッチにおいて可能になり、インターネット等のＷＡＮを介してデー

タをターゲットへエクスポート（又はインポート）することができる。

【０１５９】

しかしながら、低速リンク上のミラーリングとローカルミラーリングとの間の１つの重要な相違点は、リモートターゲットとの通信における固有の待ち時間である。例えば、ＷＡＮ上でリモートターゲットと通信する場合の平均待ち時間は８μｓ／マイルである。従って、リモートターゲットが地球の裏側にある場合、待ち時間は１００ｍｓ（往復２００ｍｓ）であり、ローカルターゲットと通信する場合よりも非常に低速であろう。

【０１６０】

１つの実施形態において、２つの（又はそれ以上の）ローカル仮想ターゲットをミラーリングする場合、前述のように、コマンドの送信後に、本発明の実施形態によるスイッチは、イニシエータ（例えば、サーバ）から書込みデータを要求する前に、全てのターゲットからのＲ２Ｔ又はＸＦＲ＿ＲＤＹを受信するのを待つことになる。その後、書込みデータは、全てのターゲットにマルチキャストされる。しかしながら、低速リンク上でのミラーリングに関しては、長いネットワーク待ち時間を避けるために、スイッチは、リモートターゲットからＲ２Ｔ又はＸＦＲ＿ＲＤＹを受信するのを待たない。その代わりに、スイッチは、ローカルターゲットからＲ２Ｔ又はＸＦＲ＿ＲＤＹを受信すると、直ちに書込みデータをイニシエータから要求してローカルターゲットに書込む。リモート装置に接続するラインカードは、リモートターゲットからＲ２Ｔ又はＸＦＲ＿ＲＤＹを受信すると、ローカルターゲットからデータを読み出し、次に、データをリモートターゲットに書込む。

【０１６１】

詳細には、図２５を参照すると、スイッチは、書込みコマンドをサーバから受信することになる（ステップ２５０２）。ローカルミラーリングの場合と同様に、入口側ＰＰＵは、コマンドをローカルターゲット及びリモートターゲットの両方の出口側ラインカードにマルチキャストすることになる（ステップ２５０４）。しかしながら、リモートターゲット宛のコマンドのフローＩＤは特別なフローＩＤなので、パケットは、他の環境で行われるように、ＰＰＵによって直接処理せ

ずに、出口側ラインカードＣＰＵに導かれることになる。ローカルターゲット宛のパケットは依然としてＰＰＵによって処理される。次に、コマンドは、それぞれの出口側ラインカードによって、各々のターゲットに、即ちローカルターゲット及びリモートターゲットに送られる（ステップ２５０６）。

【０１６２】

ネットワーク待ち時間に起因して、Ｒ２Ｔ又はＸＦＲ＿ＲＤＹは、最初に、ローカルターゲットからスイッチによって受信されることになる（ステップ２５０８）。次に、Ｒ２Ｔ又はＸＦＲ＿ＲＤＹは、イニシエータ（サーバ）に返送されることになる（ステップ２５１０）。次に、イニシエータは、自身の書込みデータをスイッチに送信し、次に、データは、書込みのためにローカルターゲットに送られることになる（ステップ２５１２）。ローカルターゲットでの書込みが終了すると、ローカルターゲットは、タスクが完了したことを示す応答パケットを送信することになる（ステップ２５１４）。

【０１６３】

最終的には、Ｒ２Ｔ又はＸＦＲ＿ＲＤＹは、ラインカードによってリモートターゲットから受信される（ステップ２５１６）。リモートターゲットに接続するラインカードのＣＰＵは書込みコマンドを送信したので、リモートＲ２Ｔ又はＸＦＲ＿ＲＤＹは同様にラインカードＣＰＵによって受信され、ラインカードＣＰＵは、リモートターゲットへのコマンドを管理する点に留意されたい。リモートターゲットに関するラインカードＣＰＵは、予め書込まれたデータを読取るために、受信したＲ２Ｔ又はＸＦＲ＿ＲＤＹをローカルターゲットに対する読み出しコマンドに変換する（ステップ２５１８）。ローカルターゲットから受信した読み出しデータは、リモートターゲットに関するラインカードのＰＰＵによって受信される（ステップ２５２０）。次に、ＰＰＵは、読み出しデータを書込みデータとしてリモートターゲットへ送る（ステップ２５２２）。書込みが完了すると、リモートターゲットは、リモートターゲットに関するラインカードＣＰＵによってどのパケットが受信されたかを示す応答パケットを送信することになる（ステップ２５２４）。ラインカードＣＰＵは、読み出しコマンド及び書込みコマンドの両方に関する状態信号を受信する。

【０１６４】

ローカル書込みが完了する前にリモートターゲットのＲ２Ｔ又はＸＦＲ＿ＲＤＹを受信した場合、リモートラインカードは、１つの実施形態において、ローカルターゲットからのデータの読み出し処理を行う前に、ローカル書込みが完了するまで待機する。

【０１６５】

読み出し又は書込みのいずれかにエラーが発生した場合、ラインカードＣＰＵは、エラーをＳＣＣに報告する。エラー発生の場合、リモートターゲットは、ローカルターゲット及びラインカードに対して非同期になる。

【０１６６】

従って、ローカルターゲットに関して、書込みコマンドは、ローカルターゲットのラインカードのＰＰＵで実行される。しかし、リモートターゲットに関して、書込みコマンドは、そのラインカードのＰＰＵが読み出しデータを書込みデータとして送る以外は、リモートターゲットのラインカードのＣＰＵによって管理される。

【０１６７】

スナップショット：

「スナップショット」とは、一般に、特定の時点まで仮想ターゲットをミラーリングし、その後、ミラーリングされたメンバを切断することにより、切断時点でミラーリングされたメンバのミラーデータをフリーズ（freeze）することである。換言すると、特定の時点でのデータの表面上の「スナップショット」が保持される。スナップショットが取得されると、ユーザは、リストアを必要とすることなく、（別の仮想ターゲットとしての）取り除かれたメンバにアクセスして、いつでも古い情報を検索することができる。従って、「スナップショット」を利用することで、本発明によるスイッチの一部のユーザは、従来のバックアップ作業及びリストア作業を行なう必要がなくなるであろう。更に、本発明によるスイッチを使用することによって、スナップショットを素早く行うことができ、所要時間は、仮想ターゲットをテープ媒体にコピーするために何時間ものバックアップウィンドウを必要とする（、及び通常はコピーされたデータへのアクセスも防

止する）場合がある従来のバックアップに比較して、わずか数ミリ秒である。また、仮想ターゲットのスナップショットは、一定の時間間隔で行うことができる。更に、各々のスナップショットは、ミラーリングされた仮想ターゲットの異なるメンバとすることができ、複数のスナップショット（例えば、火曜日のスナップショット、水曜日のスナップショット等）の利用可能性を最適化するものである。

【０１６８】

特に、図２６を参照すると、本発明の１つの実施形態によるスナップショットサービスを行うために、スナップショット要求は、管理ステーションからスイッチによって受信される（ステップ２６０２）。ＳＣＣは、入口側ラインカードＣＰＵ（サーバに接続するラインカード）に、ミラーリングされたメンバを取り除くための変更を知らせる（ステップ２６０４）。また、ＳＣＣは、ＳＣＣデータベースの仮想ターゲットオブジェクトを更新する。ラインカードＣＰＵは、もはや取り除かれたメンバを反映しないように仮想ターゲットの（ＰＰＵ　ＳＲＡＭ内にある）ＶＴＤに記憶されたフローＩＤを更新する（ステップ２６０６）。この変更により、受信した書込みデータは、取り除かれたメンバに対してマルチキャストされない。ＶＴＤが更新されると、ＣＰＵは、ＳＣＣに対する変更を了解し、ＳＣＣは、スナップショットが完了したことを示す応答信号を管理ステーションに返送する（ステップ２６０８）。

【０１６９】

更に、任意のスナップショットを開始する前に、仮想ターゲットに対する未処理要求があってはいけない。従って、スナップショットが行われる場合、１つの実施形態において、仮想ターゲットに対する全ての未処理要求を休止するようにサーバに通知する必要がある。サーバの作動は、スナップショット後に再開される。

【０１７０】

仮想ターゲットクローニング（複製）：

本発明によるスイッチは、ミラーリングされた仮想ターゲットへの新しいメンバの追加をサポートすることができ、本明細書ではクローニング（又は複製）と呼

び、オフラインで仮想ターゲットを取得することなく行なうことができる。一般に、新メンバは、ＳＣＣデータベースの仮想ターゲットオブジェクトを変更することによって追加され、ミラーリングされたターゲットの内容は、新メンバに複製されるが、仮想ターゲットに対する通常のアクセスは依然として有効である。仮想ターゲットのサイズにもよるが、複製を完了するにはある程度の時間を必要とするはずである。しかしながら、複製は、スイッチによって制御され、ユーザには見えず、一般にサーバによる仮想ターゲットへのアクセスを妨害しない。

【０１７１】

詳細には、図２７を参照すると、複製要求はＳＣＣによって受信される（ステップ２７０２）。ＳＣＣは、クローニング継続中のフラグを仮想ターゲットオブジェクトに設定し（ステップ２７０４）、サーバに接続するラインカードのＣＰＵに変更を知らせる（ステップ２７０６）。ラインカードＣＰＵは、ＰＰＵ　ＳＲＡＭ内のＶＴＤを更新し、仮想ターゲットのフローＩＤを変更して新メンバを追加する（ステップ２７０８）。フローＩＤの変更により、受信書込みデータはこの時点でマルチキャストされる。しかしながら、受信書込みはマルチキャストされるが、フローＩＤは新メンバに関する出口側ラインカードＣＰＵへ書込みデータを導くように設定され、その結果、ＰＰＵのかわりにＣＰＵが書込みデータを処理する。以下に詳細に説明するように、出口側ラインカードＣＰＵは、複製が完了するまで新メンバに対するトラフィックを一時的に管理することになる。

【０１７２】

新メンバに接続するラインカードのＣＰＵは、新メンバにコピーされることになる仮想ターゲットの内容を指定する変更記述子を作成する（ステップ２７１０）。記述子は、オフセット及びブロックカウント（オフセット、ブロックカウント）を示す。例えば、１０ＧＢターゲットをコピーするための変更記述子は（０，２０，０００，０００）であり、１つの実施形態において、各々のブロックは５１２バイトであり、１０ＧＢターゲットは２０００万個のブロックを有する点に留意されたい。ラインカードＣＰＵは、変更記述子を使用して、１度に数ブロックのコピー機能を管理する。最初に、ラインカードＣＰＵは、書込みコマンドを新メンバに送信する（ステップ２７１２）。Ｒ２Ｔ又はＸＦＲ＿ＲＤＹが

返送されると（ステップ２７１４）、ラインカードＣＰＵは、旧メンバに対する読み出し要求を初期化するが、読み出しデータを新メンバのラインカードＣＰＵに導くフローＩＤを指定する（ステップ２７１６）。何らかの読み出し又は書込みエラーが発生すると、コピーは異常終了してＳＣＣに報告される。

【０１７３】

変更記述子はブロック集合のコピー後に更新される（ステップ２７１８）。例えば、５０個のブロックのコピー後に、前述の変更記述子は、最初の５０個のブロックがこの時点では同期状態にないので（５０，１９，９９９，９５０）になる。ブロック集合のコピー処理は、全ブロックがコピーされるまで続く（ステップ２７２０）。

【０１７４】

仮想ターゲットが複数の領域で構成され、各々の領域が異なるラインカードを介してスイッチに結合されている場合、両方の領域に関する複製処理を同時に実行することができる。しかし、両方の領域が同じラインカードを介してスイッチに結合されている場合、複製プロセスは、順次実行する必要がある。即ち、第１の領域の複製が完了するまで、第２の領域は複製できない。

【０１７５】

一時的に、複製処理の間、仮想ターゲットへの書込み要求は、サーバから受信することができるとともに全てのミラーリングされたメンバに書込む必要があり、受信処理において仮想ターゲットの全データであるメンバを含む。この場合、書込み要求がマルチキャストされると、その書込み要求はミラーリングされたターゲットの旧メンバと同様に、それぞれのライン カード上のＰＰＵによって処理されるのではなく、新メンバのラインカードのＣＰＵによって受信される（ステップ２７２２）。ラインカードＣＰＵは、書込み位置を変更記述子のオフセットと照らし合わせることによって、書込みが未コピーのブロックのいずれかに対するものであるか否かを判定する（ステップ２７２４）。書込みがコピー済みのデータブロックに対するものである場合、書込みコマンドは、単純にＰＰＵに送られる（ステップ２７２６）。しかしながら、書込みが未コピーのデータブロックに対するものである場合、新メンバに対する書込みは破棄され（ステップ２７２

８）、そしてイニシエータに対してタスク完了の応答信号を送信する。それであってもなお、新しいデータは、最終的には続行している複製処理中に旧メンバから新メンバへコピーされることになる。この処理は、完了するまで複製を実行し続ける（ステップ２７２０）。

【０１７６】

別の方法において、複製処理中に仮想ターゲットに対する書込み要求が受信されると、仮想ターゲットに対して行われた変更は、ラインカードＣＰＵによって追跡することができる。複製が完了すると、その変更及び追跡部分を更新できる。

【０１７７】

複製処理が完了すると、ラインカードＣＰＵは、ＳＣＣに通知する（ステップ２７３０）。ＳＣＣは、クローニング継続中のフラグを解除するために仮想ターゲットオブジェクトを更新する（ステップ２７３２）。イニシエータに接続する入口側ラインカード上ではフローＩＤが更新されるので、書込みコマンドは、新メンバのラインカードＣＰＵに導かれるのではなく、通常どおりＰＰＵへ進む（ステップ２７３４）。

【０１７８】

第三者コピー：

第三者機能は、オフライン仮想ターゲット（アクセスされていないもの）を書込み可能ＣＤ又はテープドライブ等のアーカイブ装置へ又はアーカイブ装置からコピーする。このコピーは、コピーが完了するまでサーバは関与せず、むしろスイッチによって実行されるので「第三者コピー」と呼ばれる。多くの実施形態において、このような第三者コピーは、予め取得した仮想ターゲットのスナップショットから行われることになる。従来システムの大部分においては、このようなコピーを行うために、ターゲット装置はスマートテープ装置等の「スマート」な装置である必要があり、このことは、装置が全体的にコピー処理に積極的に関与し、且つ少なくとも部分的にコピー処理を制御することを意味する。対照的に、本システムの第三者コピーサービスは、記憶装置スイッチ外部の処理能力によるものではない。

【０１７９】

図２８を参照すると、スイッチは、コピー要求を管理ステーションから受信することになる（ステップ２８０２）。ＳＣＣは、仮想ターゲットへの書込みに関する未処理接続が確実に存在しないようにする（ステップ２８０４）。コピー時に、仮想ターゲットは、１つの実施形態において、読み出しにのみ利用可能である。次に、ＳＣＣは、ＳＣＣデータベースの仮想ターゲットオブジェクトにコピー継続中のフラグを設定して、ターゲットへの書込みに関する他の接続が確実に存在しないようにする。次に、ＳＣＣは、コピー宛先装置に接続されたラインカードのＣＰＵにコピーを実行するよう指示する（ステップ２８０８）。

【０１８０】

各々の仮想ターゲットは複数の領域から構成することができ、各々の領域は異なる物理的装置上にあってもよい。従って、宛先ラインカードのＣＰＵは、各々の領域からデータを取得する必要がある。これを行うために、宛先ラインカードのＣＰＵは、各々の領域の各々のラインカードに領域記述子を送信する（ステップ２８１０）。領域記述子は、領域及び宛先ラインカード（宛先コピー用）を指定する。次に、それぞれの領域に関するラインカードの各々のＣＰＵは、それぞれのＰＰＵ（例えば、ＶＴＤ及びＣＡＭ）をセットアップして、ＰＰＵが読み出し要求を処理できるようにする（ステップ２８１２）。

【０１８１】

領域ラインカードがセットアップ状態になると、次に、宛先ラインカードＣＰＵは、書込みコマンドを宛先装置に送信する（ステップ２８１４）。Ｒ２Ｔ又はＸＦＲ＿ＲＤＹを宛先ラインカードによって宛先装置から受信すると（ステップ２８１６）、宛先ラインカードは、それぞれの領域ラインカードを経由して読み出しコマンドを領域の１つに送る（ステップ２８１８）。読み出しデータは、宛先ラインカードに直接送信され、宛先ラインカードＰＰＵによって書込みデータとして処理され（ステップ２８２０）、書込みデータは、宛先装置に書込まれる。この処理は、領域全体がコピーされるまで繰り返される。何らかのエラーが発生するとコピーは終了する。次に、全ての領域がコピーされていない場合（ステップ２８２２）、この処理はステップ２８１４に戻り、次の領域のコピーが実行される。全ての領域がコピーされた場合（ステップ２８２２）、宛先ラインカード

のＣＰＵは、コピー完了をＳＣＣに報告する（ステップ２８２４）。エラー完了の場合、ＳＣＣは、コピーを終了する。しかし、コピーがエラーなしで完了した場合、ＳＣＣは、ＳＣＣデータベースの仮想ターゲットオブジェクトのコピー継続中のフラグをリセットし（ステップ２８２６）、管理ステーションに完了状態を報告する（ステップ２８２８）。ソース仮想ターゲットは、この時点で再び書込みができるようになる。

【０１８２】

定期的スナップショット及びバックアップ：

本発明の実施形態によるスイッチは、仮想ターゲットの定期的スナップショット及びバックアップを行うことができる。このようなバックアップ機能は、一般に３つのステップを含む。

１．仮想ターゲットをスナップショットする。

２．スナップショットから仮想ターゲットを第三者コピーする。

３．仮想ターゲットへスナップショットを送るメンバをミラーリングされたメンバとして再加入させて、最新のミラーリングされたデータの全てをそのメンバに持ち込む。

【０１８３】

第３のステップは、（前述の）複製によって、又は、スナップショットが取得された時間からメンバが再加入されるまで、仮想ターゲットの更新データを追跡する他の方法によって実行することができる。例えば、仮想ターゲットに対して行われた全ての変更記録を保持することができ、そして次に、ミラーリングされたメンバは、ミラーリングされたメンバとして仮想ターゲットを再加入させた時点で、単純にこれらの変更内容で更新される。

【０１８４】

ユーザが多数の記憶空間を有する場合、ユーザは各々のスナップショット仮想ターゲットにアクセスできるはずなので、第２のステップ及び第３のステップは必要ない場合もある。従って、このことはスナップショットターゲットを割り当て、且つネーミングを行うという問題に過ぎない。例えば、今週は就業日毎に、過去半年は月別に、その後は、四半期別に仮想ターゲットをバックアップするこ

とになっている場合、有限のスナップショットターゲット集合のみを割り当てる必要があり、以下のように命名できる。

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｍａｓｔｅｒ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｍｏｎｄａｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｔｕｅｓｄａｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｗｅｄｎｅｓｄａｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｔｈｕｒｄａｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｆｒｉｄａｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｆｅｂｒｕａｒｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｍａｒｃｈ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ａｐｒｉｌ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｍａｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｊｕｎｅ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．ｊｕｌｙ

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．２０００ｑ３

ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖ

ｅｒ．ｂａｃｋｕｐ．２０００ｑ４
ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．２００１ｑ１
ｉｑｎ．ｃｏｍ．ｍａｒａｎｔｉｎｅｔｗｏｒｋｓ．ｃｏｍｐａｎｙ．ｓｅｒｖｅｒ．ｂａｃｋｕｐ．２００１ｑ２

【０１８５】

スイッチは、スナップショットターゲットを割り当て、且つ既知のポリシーに基づいて定期的な作動のスケジュールを組む。また、スイッチは、ターゲットのネーミング（命名）及びリネーミング（改名）を管理する。例えば、バックアップ２００１ｑ３については、スイッチは、backup．２０００ｑ３のターゲットを再使用して、backup．２００１．ｑ３のためにネーミングを行うことになる。

【０１８６】

リストア：

種々の理由から、多くの業界では、アーカイブ媒体（例えば、テープ又はＣＤ等の一般的に取り外し可能な媒体又は携帯式媒体）にデータのバックアップコピーを保存する必要がある。スイッチは、第三者コピー機能を利用して、バックアップ又はスナップショットターゲットをアーカイブ媒体に移すことができる。スイッチは、アーカイブ媒体をデータベース上で追跡する。アーカイブ媒体のコピーが行われる度に、ＳＣＣは、全ての宛先領域を判定するために仮想ターゲットオブジェクトを取り込み、そして、記録は、媒体を追跡するために管理ステーションでデータベースに入力される。管理ステーションを使用して、ユーザは、テープ又はＣＤ集合のアーカイブ媒体のリストを閲覧し、リストアのために１つを選択することができる。

【０１８７】

リストア操作自体は、スイッチによってスケジュールされる別の第三者コピー機能である。しかしながら、この操作は、誰かが媒体をテープ又はＣＤドライブに入れる必要があるのでユーザの介入を伴う。しかしながら、本明細書で説明した他の記憶装置の場合と同様に、ソースターゲット装置のＣＰＵは、リストア操作の作業を制御するが、複数の宛先ＳＰＵは１つずつ必要とされる。

【０１８８】

本発明の１つの実施形態によるスイッチは、リストア処理の３つの異なる優先順位、即ち、緊急、重要、及び通常をサポートする。緊急リストアは、システム上の現在のトラフィック状況に無関係に直ちに開始される。重要リストアは、システムが混雑している場合には開始されないが、数時間以内に開始される。通常リストアは、システムのトラフィック混雑に応じて２４時間以内に完了される。

【０１８９】

結論：

従って、本発明の実施形態に基づいて、パケットの分類、パケット上での仮想化機能の実行、パケットの任意の所要プロトコル変換の実行を含む、データパケットのワイヤ速度での処理を可能にする記憶装置スイッチが開示される。従来の方法と比較すると、開示されたアーキテクチャによって、パケットを処理するために必要な時間を最小にすることができる。このようなワイヤ速度での処理は、ある意味では、スイッチの処理能力を全てのラインカードに分散して、全体的にバッファリングの必要性を回避することによって達成される。このような分散された処理能力は、高帯域幅を有するだけでなく、拡張が容易なシステムを可能にする。更に、自身のラインカードを使用するこのようなスイッチは、サーバ不要の記憶サービス、即ち、スイッチの外部にはその運用の制御に必要な実体物が全くないサービスを行うこともできる。

【０１９０】

以上説明した特定の実施形態は、本発明の原理を例示するものに過ぎず、当業者であれば、本発明の技術範囲及び技術思想から逸脱することなく種々の変更を行うことができることを理解されたい。従って、本発明の技術範囲は、特許請求の範囲によってのみ限定される。

【図面の簡単な説明】

【０１９１】

【図１】従来のシステムによるＳＡＮの概略機能ブロック図である。

【図２】従来の方法によるプロトコル間のインタフェース処理に使用される装置の概略機能ブロック図である。

【図３】本発明の実施形態による記憶装置スイッチを使用するＳＡＮシステムの概略機能ブロック図である。

【図４】本発明の実施形態による記憶装置スイッチを使用するシステムの別の実施形態の概略機能ブロック図である。

【図５】本発明の実施形態による記憶装置スイッチを使用するシステムの更に別の実施形態の概略機能ブロック図である。

【図６】本発明の実施形態による記憶装置スイッチの概略機能ブロック図である。

【図７】本発明の実施形態による記憶装置スイッチに使用されるラインカードの概略機能ブロック図である。

【図７ａ】本発明の実施形態による記憶装置スイッチで使用される仮想ターゲット記述子の概略ブロック図である。

【図８ａ】従来公知のｉＳＣＳＩ　ＰＤＵの概略ブロック図である。

【図８ｂ】従来公知のｉＳＣＳＩ　ＰＤＵの概略ブロック図である。

【図８ｃ】従来公知のｉＳＣＳＩ　ＰＤＵの概略ブロック図である。

【図８ｄ】従来公知のｉＳＣＳＩ　ＰＤＵの概略ブロック図である。

【図８ｅ】従来公知のｉＳＣＳＩ　ＰＤＵの概略ブロック図である。

【図８ｆ】従来公知のファイバ・チャンネルプロトコル（ＦＣＰ）フレーム及びペイロードの概略ブロック図である。

【図８ｇ】従来公知のファイバ・チャンネルプロトコル（ＦＣＰ）フレーム及びペイロードの概略ブロック図である。

【図８ｈ】従来公知のファイバ・チャンネルプロトコル（ＦＣＰ）フレーム及びペイロードの概略ブロック図である。

【図８ｉ】従来公知のファイバ・チャンネルプロトコル（ＦＣＰ）フレーム及びペイロードの概略ブロック図である。

【図９ａ】本発明の実施形態による、ＰＡＣＥにおいて行われる、入口側のｉＳＣＳＩパケット分類処理を示す流れ図である。

【図９ｂ】本発明の実施形態による、ＰＡＣＥにおいて行われる、出口側のｉＳＣＳＩパケット分類処理を示す流れ図である。

【図１０ａ】本発明による記憶装置スイッチに入るときのＴＣＰパケット及びＴＣＰパケットのブロック図であり、パケットが記憶装置スイッチ内での使用に適するように変更される方法を示す。

【図１０ｂ】本発明による記憶装置スイッチに入るときのＴＣＰパケット及びＴＣＰパケットのブロック図であり、パケットが記憶装置スイッチ内での使用に適するように変更される方法を示す。

【図１１】本発明の実施形態による記憶装置スイッチで使用されるローカルヘッダの概略ブロック図である。

【図１２ａ】本発明の実施形態による、ＰＡＣＥにおいて行われる、入口側のＦＣＰフレーム分類処理を示す流れ図である。

【図１２ｂ】本発明の実施形態による、ＰＡＣＥにおいて行われる、出口側のＦＣＰフレーム分類処理を示す流れ図である。

【図１３ａ】本発明の実施形態による、ＰＰＵにおいて行われる、入口側の分類処理を示す流れ図である。

【図１３ｂ】本発明の実施形態による、ＰＰＵにおいて行われる、出口側の分類処理を示す流れ図である。

【図１４】本発明の実施形態による、コマンドパケット又はフレームに関する入口側の仮想化処理を示す流れ図である。

【図１４ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、入口側での（イニシエータサーバ／ポートからの）コマンドパケットのヘッダ及びＩＴＣＢを示す。

【図１５】本発明の実施形態による、コマンドパケット又はフレームに関する出口側の仮想化処理を示す流れ図である。

【図１５ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、出口側での（ファブリック／トラフィックマネージャからの）コマンドパケットのヘッダ及びＥＴＣＢを示す。

【図１６】本発明の実施形態による、Ｒ２Ｔ／ＸＦＲ＿ＲＤＹパケット又はフレームに関する入口側での仮想化処理を示す流れ図である。

【図１６ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ

及びＥＴＣＢ）のブロック図であり、入口側での（ターゲット記憶装置／ポートからの）Ｒ２Ｔ／ＸＦＲ＿ＲＤＹパケットに関するヘッダ及びＥＴＣＢを示す。

【図１７】本発明の実施形態による、Ｒ２Ｔ／ＸＦＲ＿ＲＤＹパケット又はフレームに関する出口側での仮想化処理を示す流れ図である。

【図１７ａ】仮想化処理中のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、出口側での（ファブリック／トラフィックマネージャからの）Ｒ２Ｔ／ＸＦＲ＿ＲＤＹパケットに関するヘッダ及びＩＴＣＢを示す。

【図１８】本発明の実施形態による、書込みデータパケット又はフレームに関する入口側での仮想化処理を示す流れ図である。

【図１８ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、入口側での（イニシエータサーバ／ポートからの）書込みデータパケットに関するヘッダ及びＩＴＣＢを示す。

【図１９】本発明の実施形態による、書込みデータパケット又はフレームに関する出口側での仮想化処理を示す流れ図である。

【図１９ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、出口側での（ファブリック／トラフィックマネージャからの）書込みデータパケットに関するヘッダ及びＥＴＣＢを示す。

【図２０】本発明の実施形態による、読み出しデータパケットに関する入口側での仮想化処理を示す流れ図である。

【図２０ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、入口側での（ターゲット記憶装置／ポートからの）書込みデータパケットに関するヘッダ及びＥＴＣＢを示す。

【図２１】本発明の実施形態による、読み出しデータパケットに関する出口側での仮想化処理を示す流れ図である。

【図２１ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、出口側での（ファブリック／トラフィックマネージャからの）書込みデータパケットに関するヘッダ及びＩＴＣＢを示す。

【図２２】本発明の実施形態による、応答パケット又はフレームに関する入口側

での仮想化処理を示す流れ図である。

【図２２ａ】仮想化処理時のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、入口側での（ターゲット記憶装置／ポートからの）応答パケットに関するヘッダ及びＥＴＣＢを示す。

【図２３】本発明の実施形態による、応答パケット又はフレームに関する出口側での仮想化処理を示す流れ図である。

【図２３ａ】仮想化処理中のローカルヘッダ及びタスク制御ブロック（ＩＴＣＢ及びＥＴＣＢ）のブロック図であり、出口側での（ファブリック／トラフィックマネージャからの）応答パケットに関するヘッダ及びＩＴＣＢを示す。

【図２４】本発明の実施形態による記憶サービスを実行するために行われるステップを示す流れ図である。

【図２５】本発明の実施形態による低速リンク上のミラーリング記憶サービスのために行われる概略的ステップを示す流れ図である。

【図２６】本発明の実施形態によるスナップショット記憶サービスのために行われるステップを示す流れ図である。

【図２７】本発明の実施形態によるクローニング記憶サービスのために行われるステップを示す流れ図である。

【図２８】本発明の実施形態による第三者コピー記憶サービスのために行われるステップを示す流れ図である。

【図１】

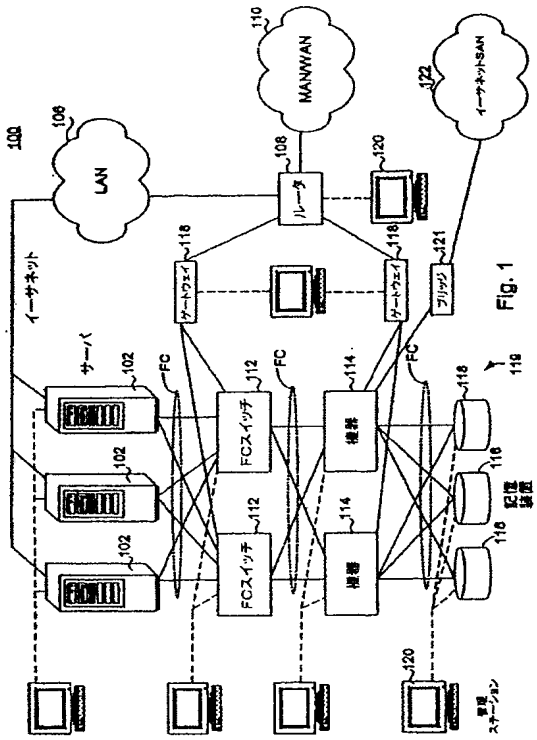


Fig. 1

【図２】

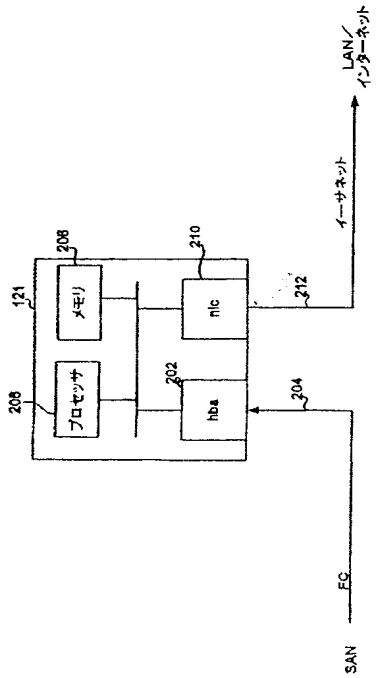


Fig. 2

【図３】

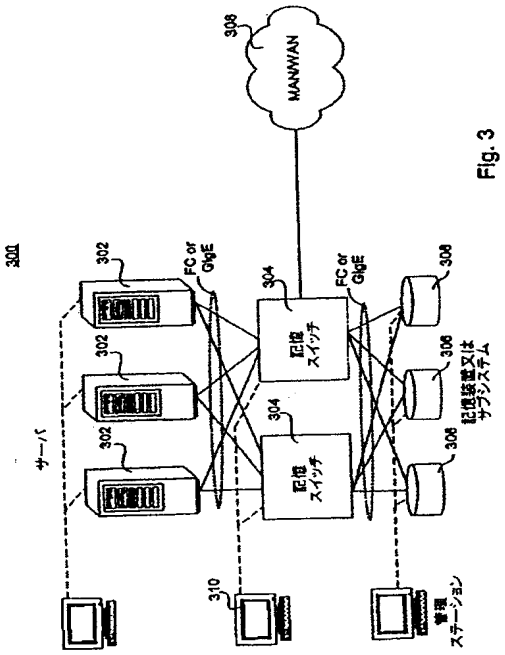


Fig. 3

【図４】

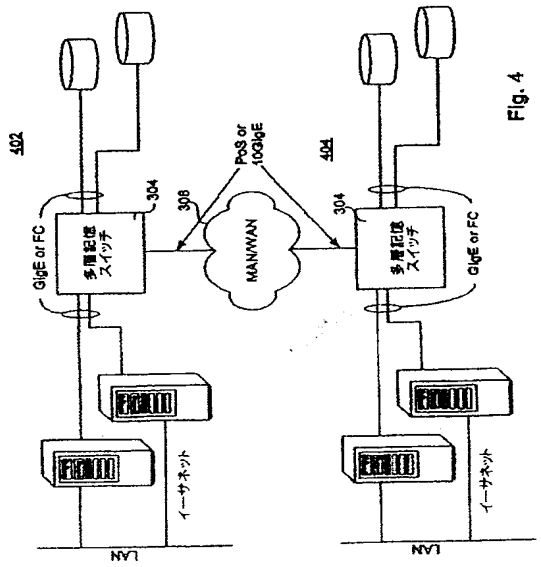


Fig. 4

【図５】

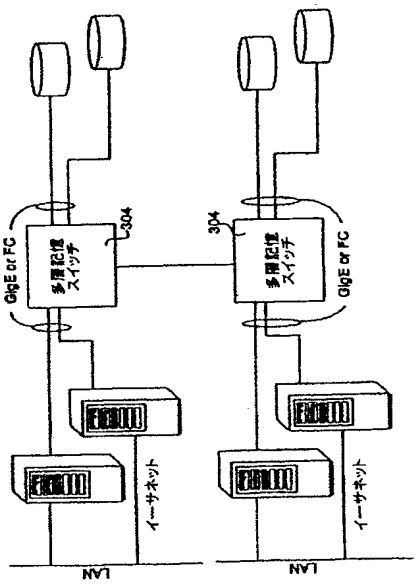


Fig. 5

【図６】

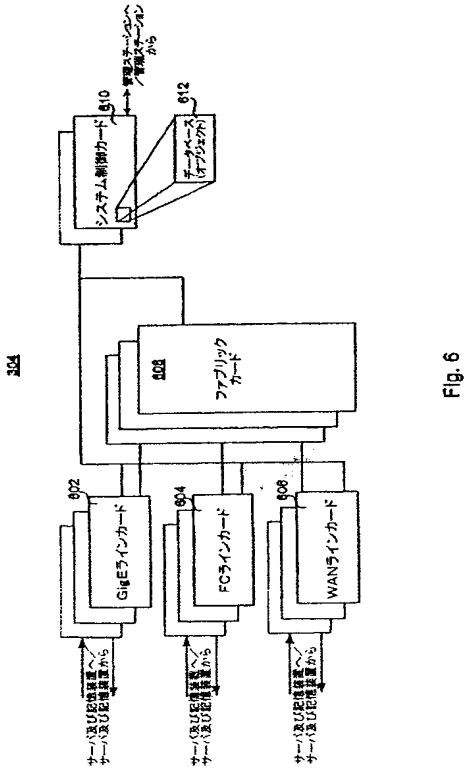


Fig. 6

【図７】

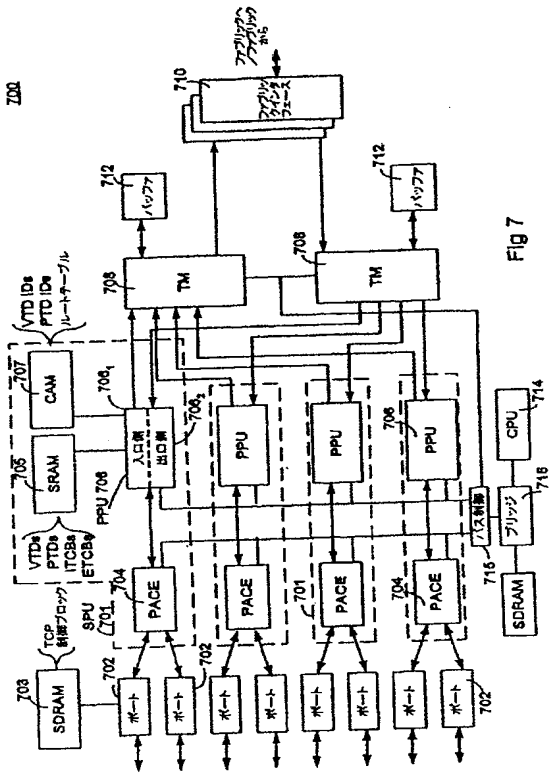


Fig 7

【図７ａ】

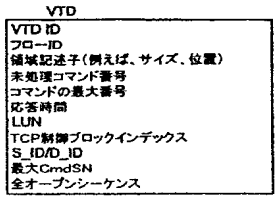
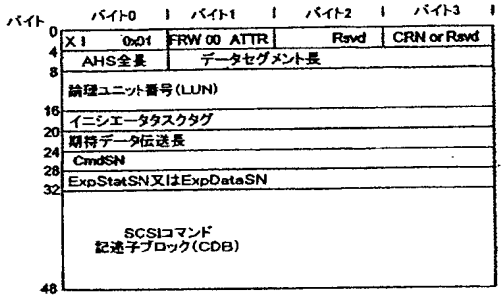


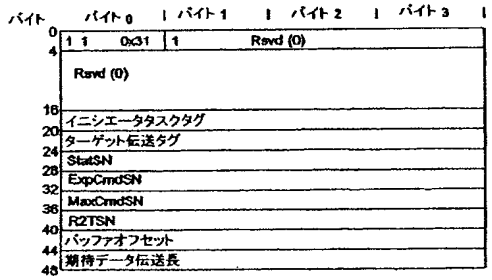
Fig. 7a

【図８ａ】



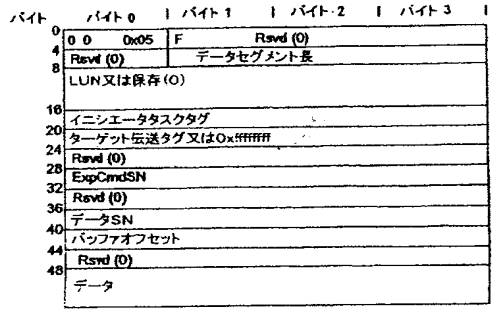
iSCSIコマンドPDU
Fig. 8a

【図８ｂ】



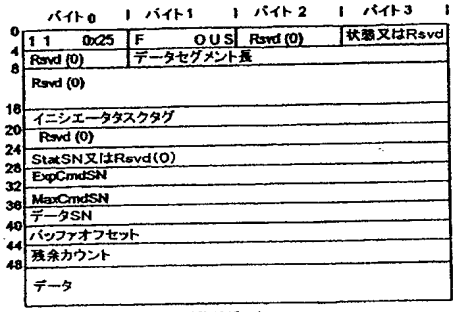
iSCSI R2T PDU
Fig. 8b

【図８ｃ】



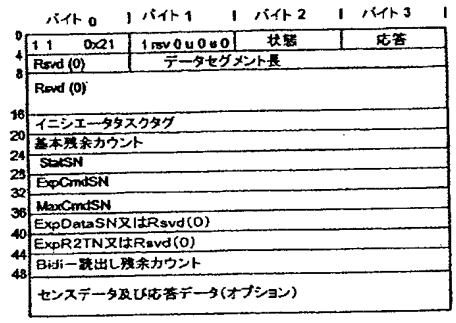
iSCSI書込みデータPDU
Fig. 8c

【図８ｄ】



iSCSI読取りデータPDU
Fig. 8d

【図８ｅ】



iSCSI応答PDU
Fig. 8e

【図８ｆ】

ビットワード	31-24	23-16	15-08	07-00
0	R_CTL	D_ID		
1	rsvd	S_ID		
2	TYPE	F_CTL		
3	SEQ_ID	DF_CTL	SEQ_CNT	
4	OX_ID		RX_ID	
5	RLTV_OFF			

FCフレームヘッダ
Fig. 8f

【図８ｇ】

フィールド名	内容	サイズ
FCP_LUN	論理ユニット番号	8バイト
FCP_CNTL	制御フィールド	4バイト
FCP_CDB	SCSIコマンド記述子ブロック	16バイト
FCP_DL	データ長	4バイト

FCP_CMNDペイロード
Fig. 8g

【図８ｈ】

フィールド名	内容	サイズ
DATA_RO	後続のFCP_DATA_IUの第1のバイトの相対オフセット	4バイト
BURST_LEN	後続のFCP_DATA_IUの長さ	4バイト
rsvd		4バイト

FCP_XFR_RDYペイロード

Fig. 8h

【図８ｉ】

フィールド名	内容	サイズ
rsvd		4バイト
rsvd		4バイト
FCP_STATUS	フィールド有効性及びSCSI状態	4バイト
FCP_RESID	残余カウント	4バイト
FCP_SNS_LEN	FCP_SNS_INFOフィールド長	4バイト
FCP_RSP_LEN	FCP_RSP_INFOフィールド長	4バイト
FCP_RSP_INFO	FCP応答情報	mバイト
FCP_SNS_INFO	FCPセンス情報	nバイト

FCP_RSPペイロード

Fig. 8i

【図９ａ】

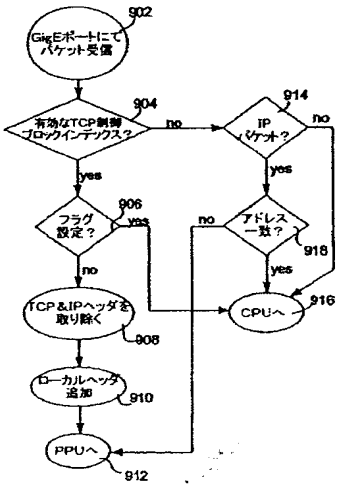


Fig 9a

（分類－PACE－iSCSi－入口側）

【図９ｂ】

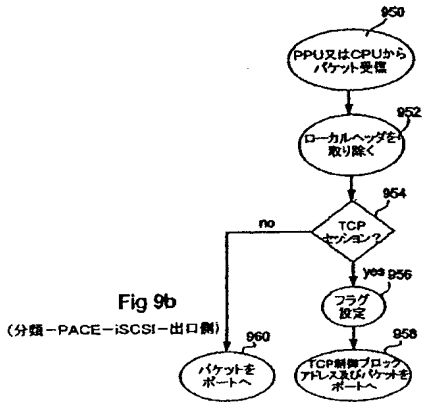


Fig 9b
(分類－PACE－iSCSI－出口側)

【図１０ａ】

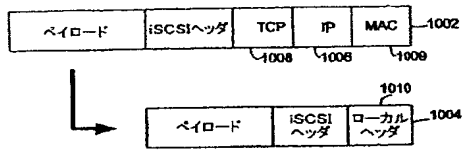


Fig. 10a

【図１０ｂ】

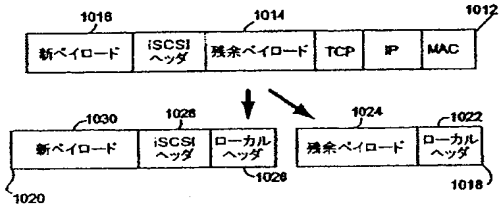


Fig. 10b

【図１１】

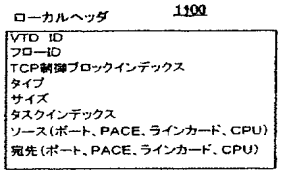


Fig. 11

【図１２ａ】

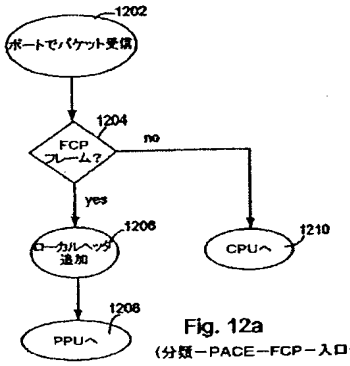


Fig. 12a
（分類－PACE－FCP－入口側）

【図１２ｂ】

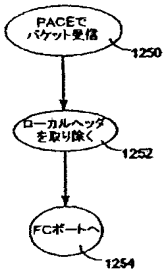


Fig. 12b
（分類－PACE－FCP－出口側）

【図１３ａ】

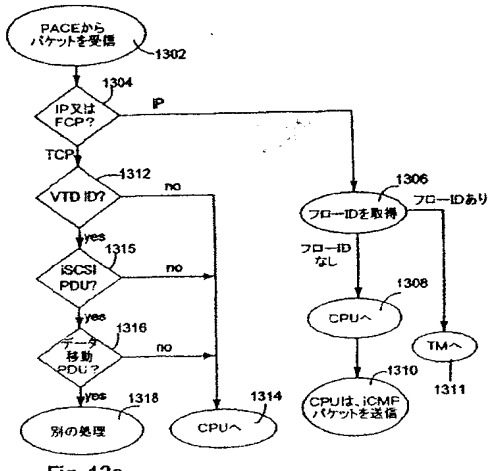


Fig. 13a
（分類－PPU－入口側）

【図１３ｂ】

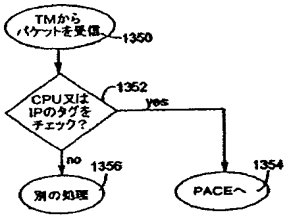


Fig. 13b
(分類－PPU－出口側)

【図１４】

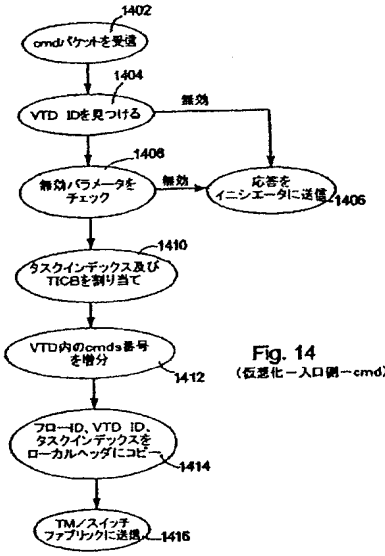


Fig. 14
(仮想化－入口側－cmd)

【図１４ａ－１５ａ】

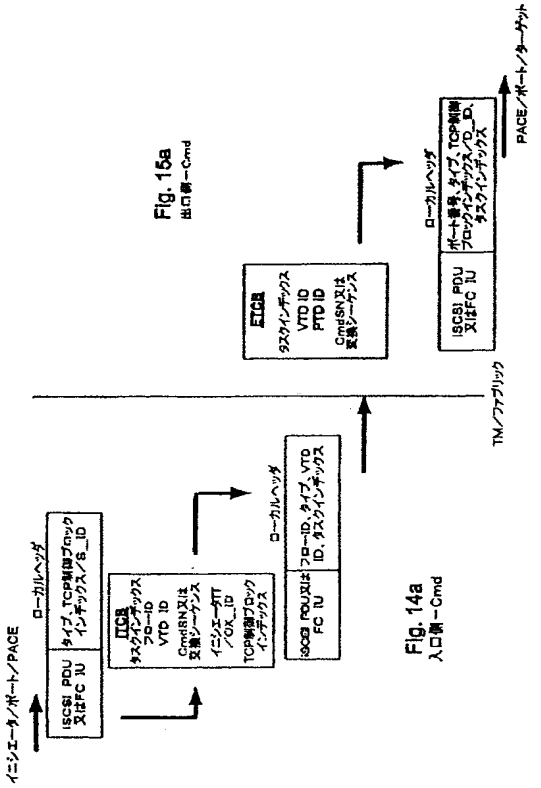


Fig. 15a
出口側－Cmd

Fig. 14a
入口側－Cmd

【図１５】

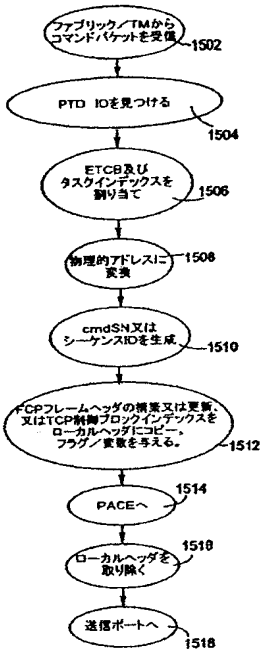


Fig. 15
(仮想化－出口側－cmd)

【図１６】

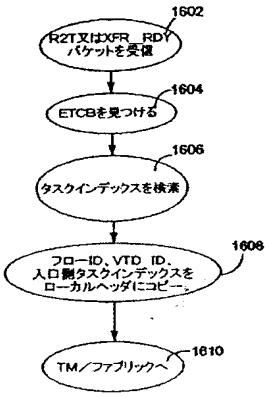


Fig. 16
(仮想化－入口側－R2T／XFR_RDY)

【図１６ａ－１７ａ】

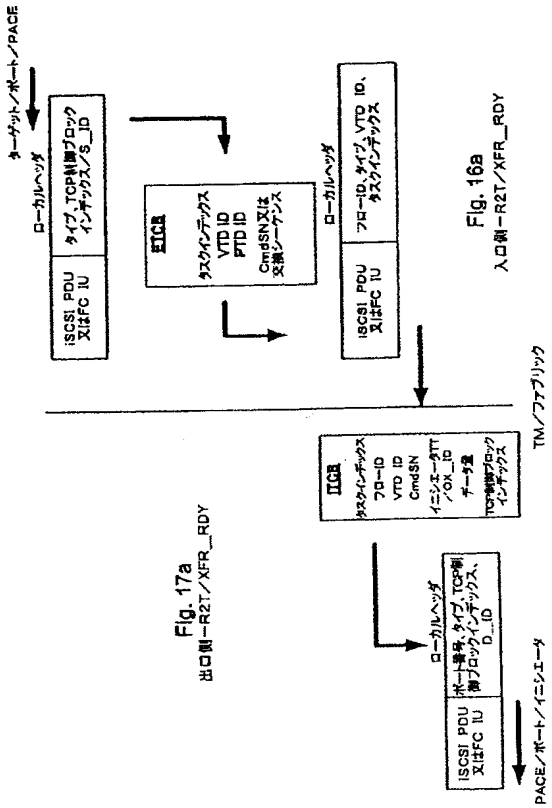


Fig. 16a
入口側－R2T／XFR_RDY

Fig. 17a
出口側－R2T／XFR_RDY

【図１７】

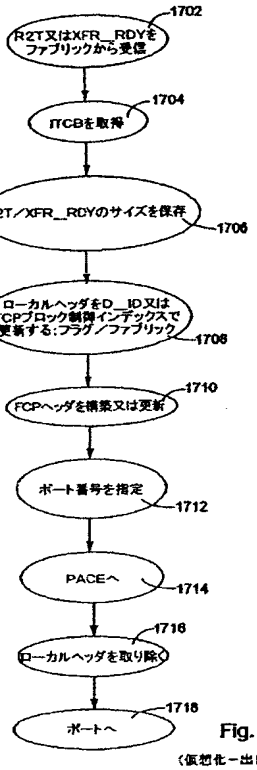


Fig. 17
(仮想化－出口側－R2T/XFR_RDY)

【図１８】

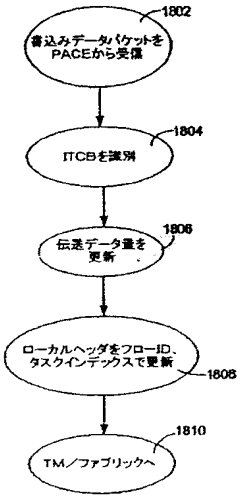


Fig. 18
(仮想化－入口側－書込みデータパケット)

【図１８ａ－１９ａ】

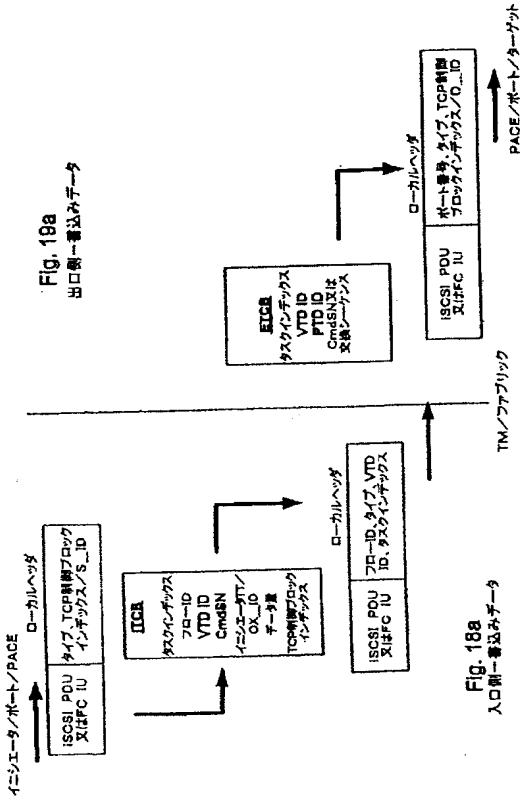


Fig. 18a
入口側－書込みデータ

【図１９】

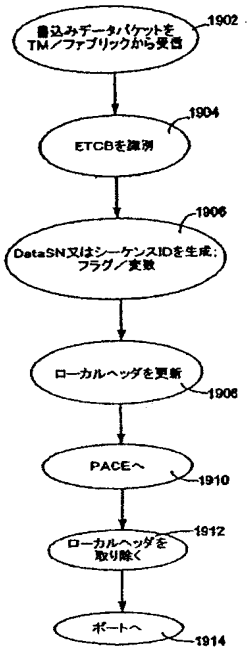


Fig. 19
（仮想化－出口側－書込みデータパケット）

【図２０】

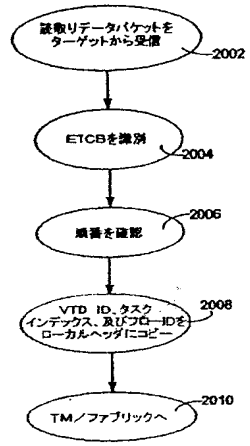
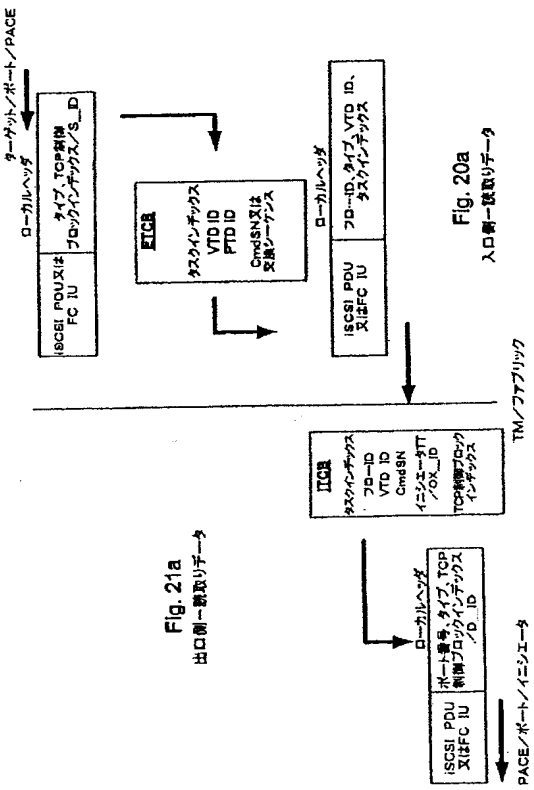


Fig. 20
（仮想化－入口側－読取りデータパケット）

【図２０ａ－２１ａ】


ターゲット／ポート／PACE
ローカルヘッダ
iSCSI PDUまたはFC IU
タイプ、TCP制御ブロックインデックス／S_ID
ETCB
タスクインデックス
VTD ID
PTD ID
CmdSNまたは交換シーケンス
ローカルヘッダ
フローID、タイプ、VTD ID、タスクインデックス
iSCSI PDUまたはFC IU
Fig. 20a
入口側－抜取りデータ
TM／ファブリック
ITCB
タスクインデックス
フローID
VTD ID
CmdSN
イニシエータTT
／OX_ID
TCP制御ブロックインデックス
Fig. 21a
出口側－抜取りデータ
ローカルヘッダ
ポート番号、タイプ、TCP制御ブロックインデックス／D_ID
iSCSI PDUまたはFC IU
PACE／ポート／イニシエータ

【図２１】

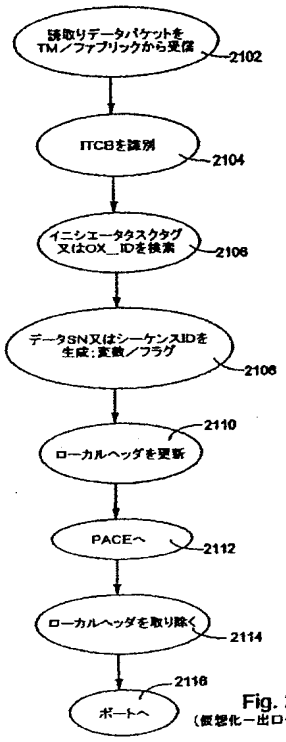


Fig. 21
（仮想化－出口側－読取りデータパケット）

【図２２】

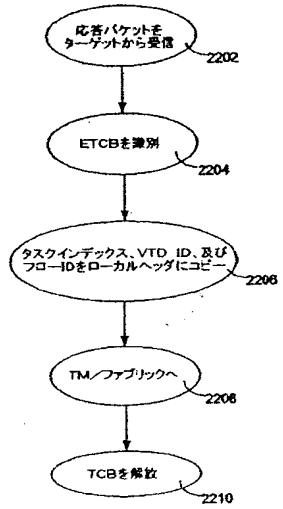


Fig. 22
（仮想化－入口側－応答パケット）

【図２２ａ－２３ａ】

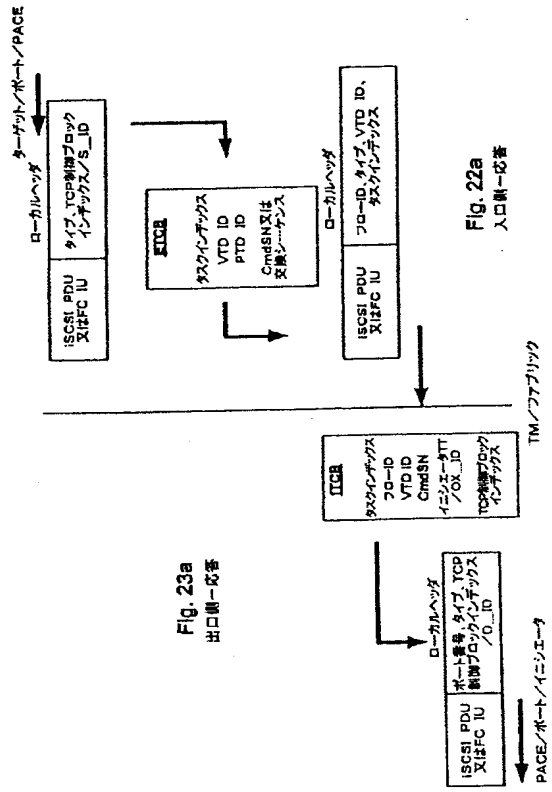


Fig. 22a
入口側一般

Fig. 23a
出口側一般

【図２３】

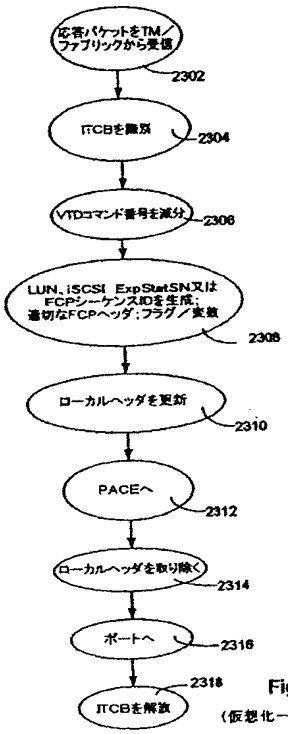


Fig. 23
（仮想化－出口側－応答パケット）

【図２４】

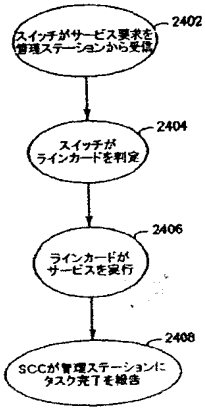


Fig.24

【図２５】

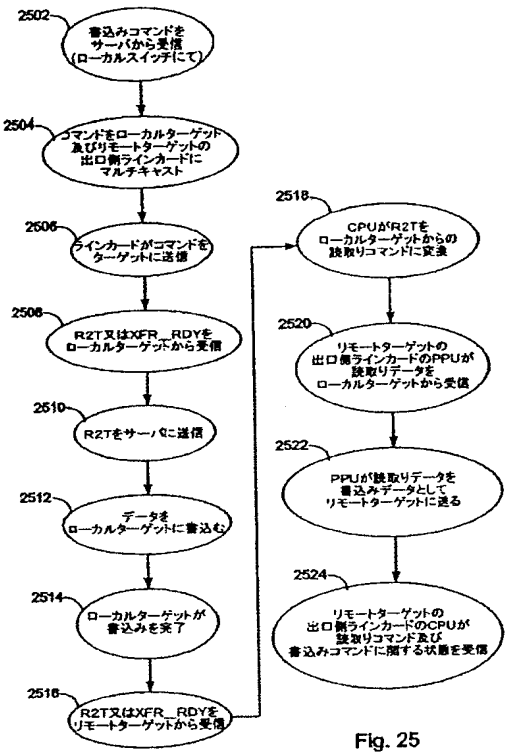


Fig. 25

【図２６】

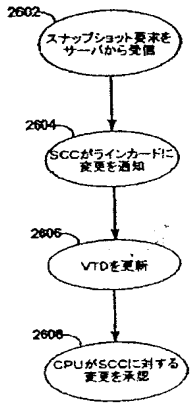


Fig. 26

【図２７】

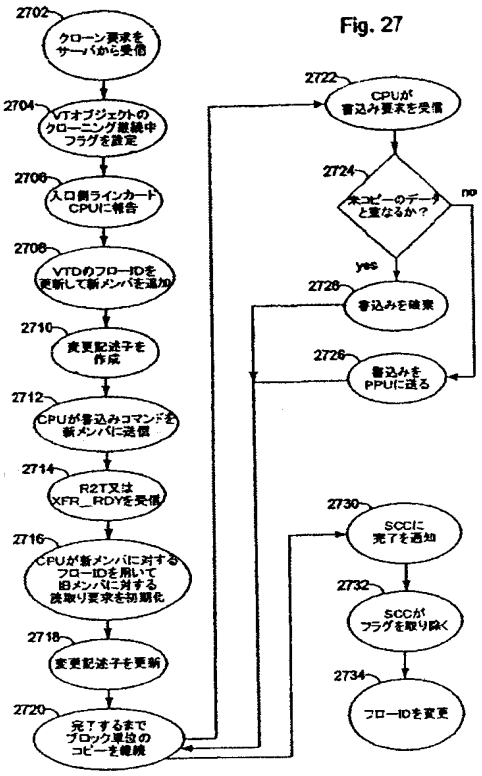


Fig. 27
2702
クローン要求をサーバから受信
2704
VTオブジェクトのクローニング継続中フラグを設定
2706
入口側ラインカードCPUに報告
2708
VTDのフローIDを更新して新メンバを追加
2710
変更記述子を作成
2712
CPUが書込みコマンドを新メンバに送信
2714
R2T又はXFR_RDYを受信
2716
CPUが新メンバに対するフローIDを用いて旧メンバに対する読取り要求を初期化
2718
変更記述子を更新
2720
完了するまでブロック単位のコピーを継続
2722
CPUが書込み要求を受信
2724
未コピーのデータと重なるか？
yes
no
2728
書込みを破棄
2726
書込みをPPUに送る
2730
SCCに完了を通知
2732
SCCがフラグを取り除く
2734
フローIDを変更

【図２８】

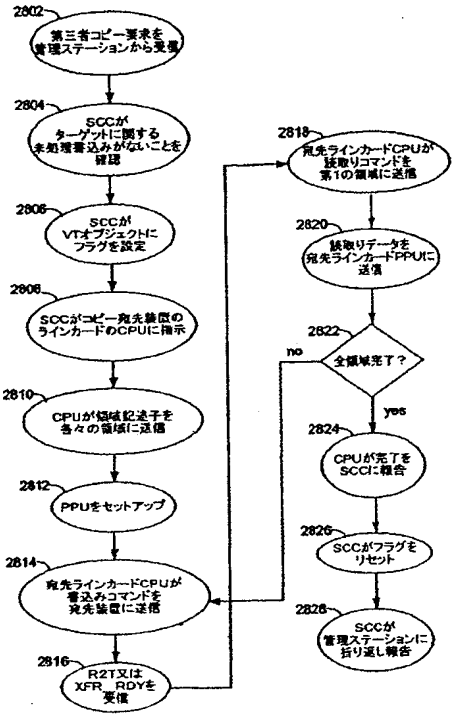

2802
第三者コピー要求を管理ステーションから受信
2804
SCCがターゲットに関する未処理書込みがないことを確認
2806
SCCがVTオブジェクトにフラグを設定
2808
SCCがコピー宛先装置のラインカードのCPUに指示
2810
CPUが領域記述子を各々の領域に送信
2812
PPUをセットアップ
2814
宛先ラインカードCPUが書込みコマンドを宛先装置に送信
2816
R2T又はXFR_RDYを受信
2818
宛先ラインカードCPUが読取りコマンドを第1の領域に送信
2820
読取りデータを宛先ラインカードPPUに送信
2822
全領域完了？
no
yes
2824
CPUが完了をSCCに報告
2826
SCCがフラグをリセット
2828
SCCが管理ステーションに折り返し報告

Fig. 28

1/38

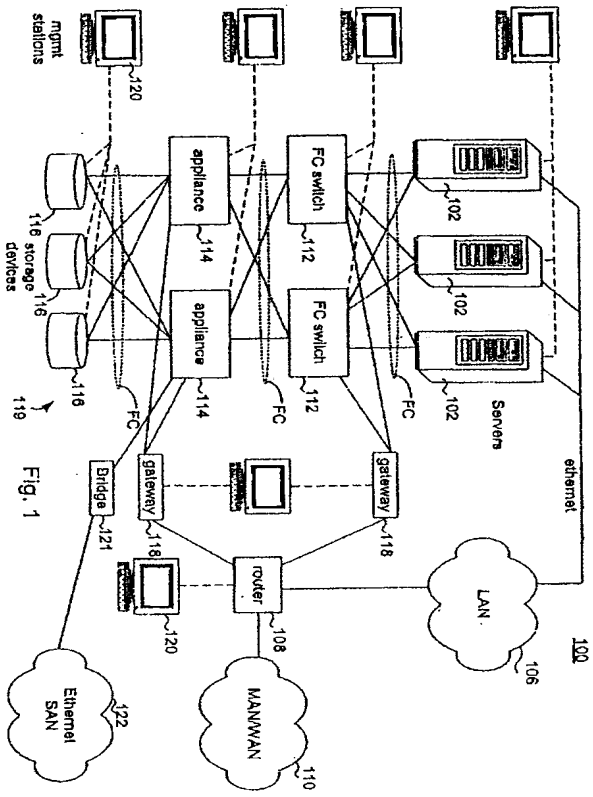

mgmt stations
120
appliance
114
FC switch
112
102
storage devices
116
FC
Servers
ethernet
Bridge
121
gateway
118
router
108
LAN
106
100
Ethernet SAN
122
MAN/WAN
110
119

Fig. 1

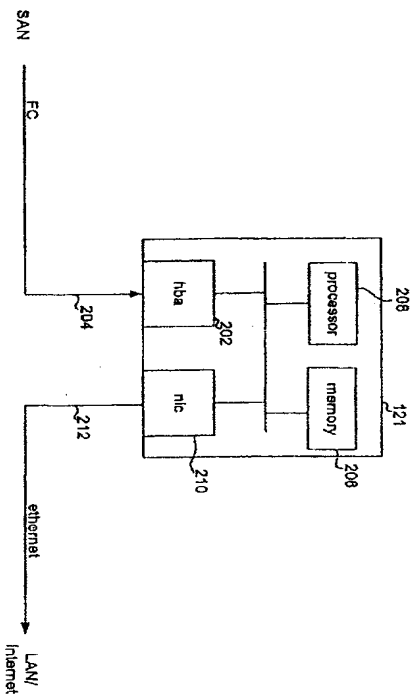

SAN
FC
204
hba
202
nic
212
210
processor
206
memory
208
121
ethernet
LAN/
Internet

Fig. 2

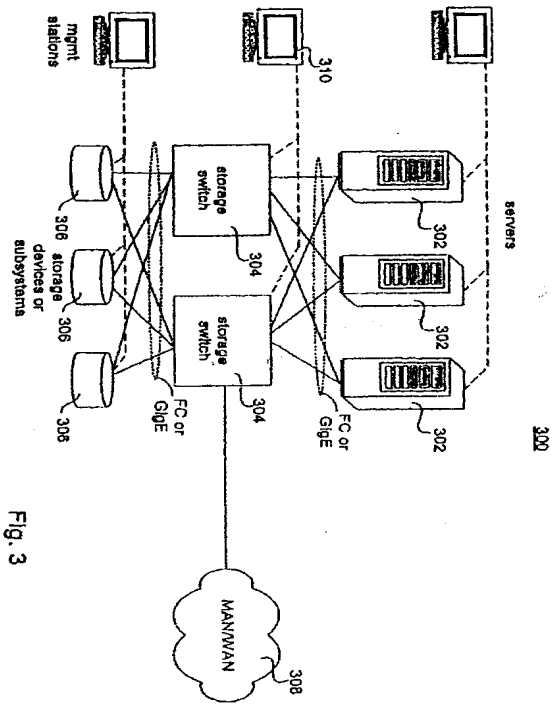

mgmt stations
310
storage switch
304
storage switch
304
servers
302
302
302
storage devices or subsystems
306
306
306
FC or GigE
FC or GigE
MAN/WAN
308
300

Fig. 3

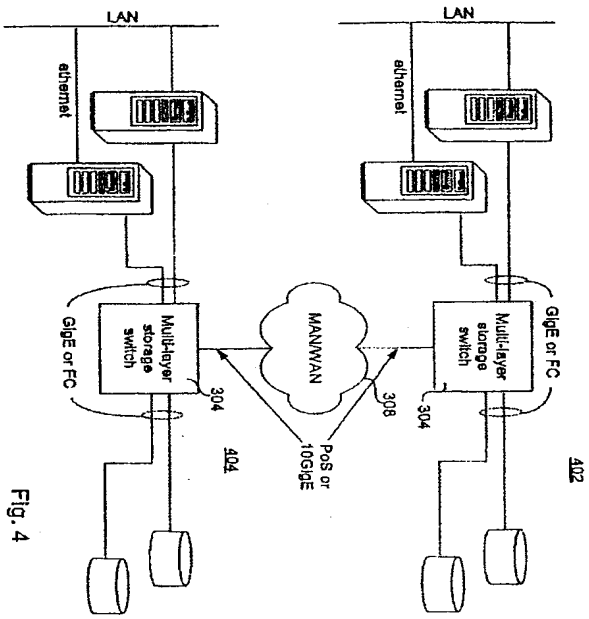

LAN
LAN
ethernet
ethernet
GigE or FC
Multi-layer storage switch
304
MAN/WAN
308
304
Multi-layer storage switch
GigE or FC
PoS or 10GigE
404
402

Fig. 4

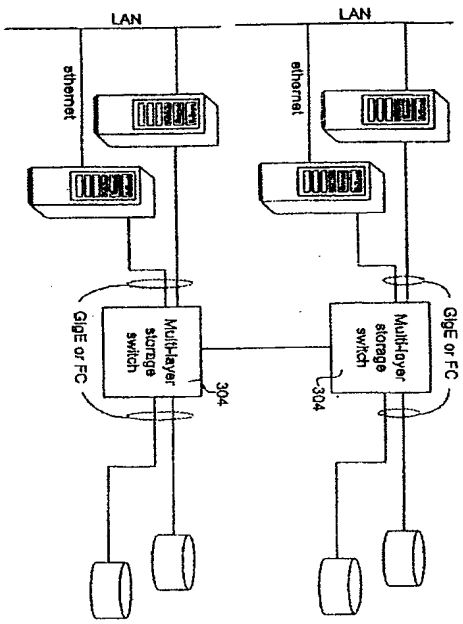

LAN
LAN
ethernet
ethernet
GigE or FC
Multi-layer storage switch
304
Multi-layer storage switch
304
GigE or FC

Fig. 5

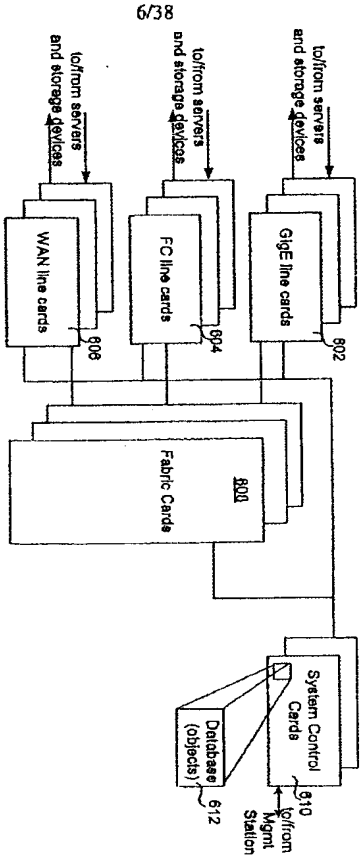

to/from servers and storage devices
to/from servers and storage devices
to/from servers and storage devices
WAN line cards
FC line cards
GigE line cards
806
804
802
Fabric Cards
801
304
System Control Cards
Database (objects)
810
812
to/from Mgmt. Station

Fig. 6

7/38

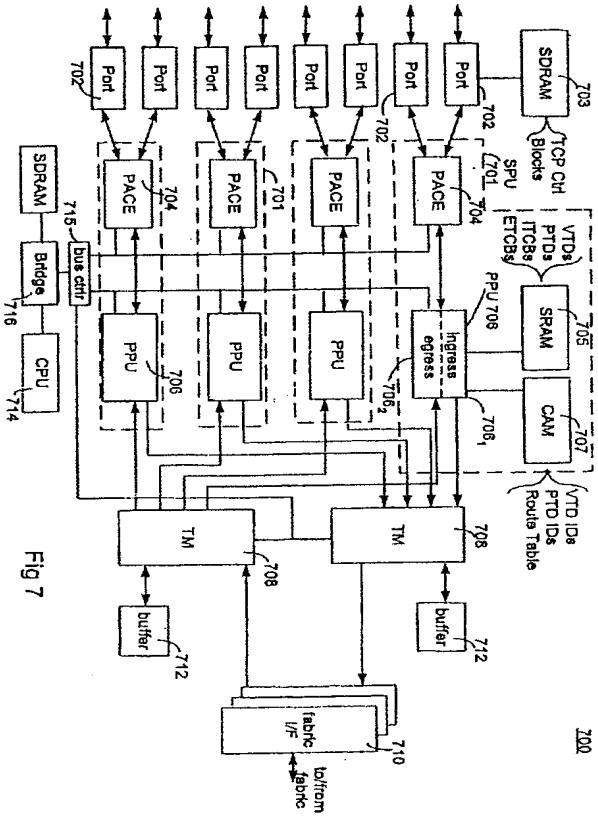

Port
SDRAM
703
TCP CH Blocks
SPU
701
PACE
704
VTDs
PTDs
ITCBs
ETCBs
PPU 708
SRAM
705
CAM
707
Ingress egress
706
PPU
TM
708
buffer
712
VTD IDs
PTD IDs
Route Table
fabric I/F
710
to/from fabric
700
SDRAM
Bridge
716
bus ctrlr
715
CPU
714

Fig 7

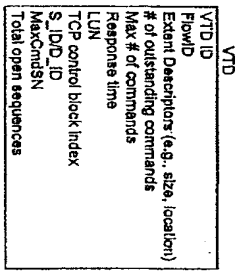
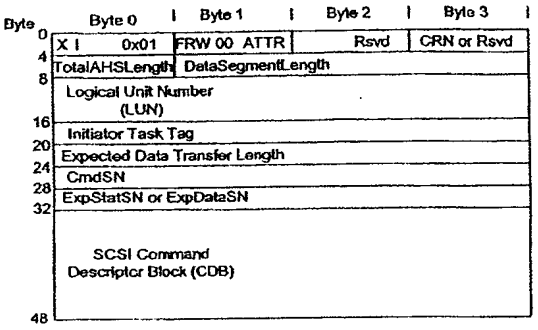
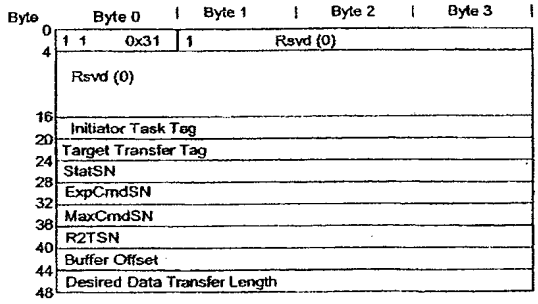

VTD
VTD ID
FlowID
Extent Descriptors (e.g., size, location)
# of outstanding commands
Max # of commands
Response time
LUN
TCP control block index
S_ID/D_ID
MaxCmdSN
Total open sequences

Fig. 7a



iSCSI Command PDU
Fig. 8a



iSCSI R2T PDU
Fig. 8b

Byte	Byte 0	Byte 1	Byte 2	Byte 3
0	0 0 0x05	F Rsvd (0)		
4	Rsvd (0)	DataSegmentLength		
8	LUN or Reserved (0)			
16	Initiator Task Tag			
20	Target Transfer Tag or 0xffffffff			
24	Rsvd (0)			
28	ExpCmdSN			
32	Rsvd (0)			
36	DataSN			
40	Buffer Offset			
44	Rsvd (0)			
48	Data			

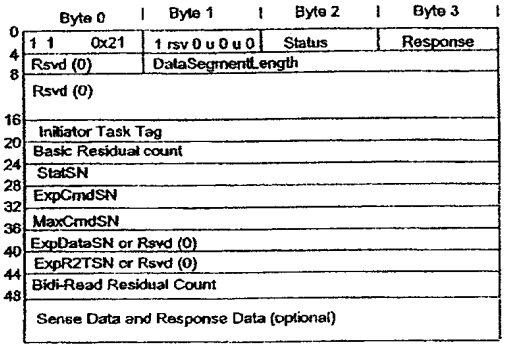
iSCSI Write Data PDU

Fig. 8c

	Byte 0	Byte 1	Byte 2	Byte 3
0	1 1 0x25	F O U S	Rsvd (0)	Status or Rsvd
4	Rsvd (0)	DataSegmentLength		
8	Rsvd (0)			
16	Initiator Task Tag			
20	Rsvd (0)			
24	StatSN or Rsvd (0)			
28	ExpCmdSN			
32	MaxCmdSN			
36	DataSN			
40	Buffer Offset			
44	Residual Count			
48	Data			

iSCSI Read Data PDU

Fig. 8d



iSCSI Response PDU

Fig. 8e

Bits Word	31-24	23-16	15-08	07-00
0	R_CTL	D_ID		
1	rsvd	S_ID		
2	TYPE	F_CTL		
3	SEQ_ID	DF_CTL	SEQ_CNT	
4	OX_ID		RX_ID	
5	RLTV_OFF			

FC Frame Header
Fig. 8f

Field Name	Description	Size
FCP_LUN	logical unit number	8 bytes
FCP_CNTL	control field	4 bytes
FCP_CDB	SCSI command descriptor block	16 bytes
FCP_DL	Data Length	4 bytes

FCP_CMND Payload
Fig. 8g

Field Name	Description	Size
DATA_RO	Relative offset of first byte of FCP_DATA IU that follows	4 bytes
BURST_LEN	length of FCP_DATA IU that follows	4 bytes
rsvd		4 bytes

FCP_XFR_RDY Payload
Fig. 8h

Field Name	Description	Size
rsvd		4 bytes
rsvd		4 bytes
FCP_STATUS	field validity and SCSI status	4 bytes
FCP_RESID	residual count	4 bytes
FCP_SNS_LEN	Length of FCP_SNS_INFO field	4 bytes
FCP_RSP_LEN	Length of FCP_RSP_INFO field	4 bytes
FCP_RSP_INFO	FCP response info	m bytes
FCP_SNS_INFO	FCP sense info	n bytes

FCP_RSP Payload
Fig. 8i

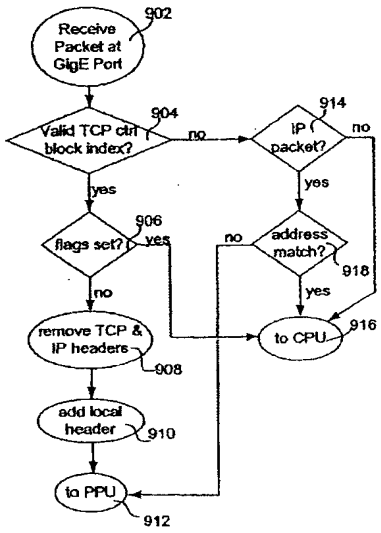


Fig 9a
(Classification - PACE - iSCSI - ingress)

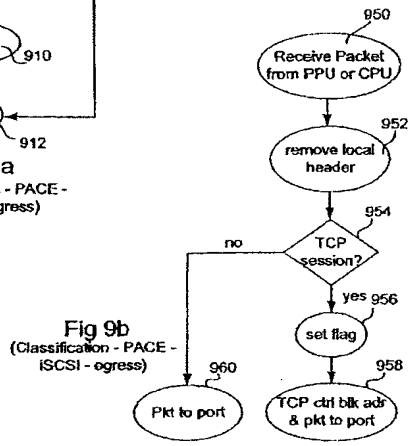


Fig 9b
(Classification - PACE - iSCSI - egress)

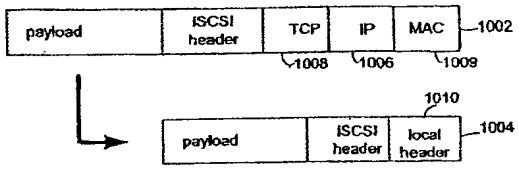


Fig. 10a

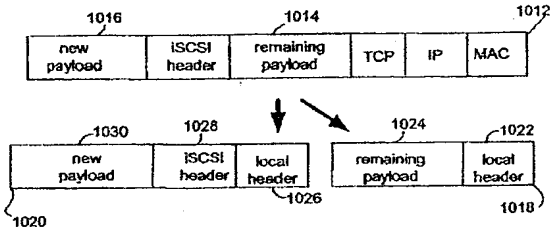


Fig. 10b

1100

Local Header

VTD ID
FlowID
TCP Control Block Index
Type
Size
Task Index
Source (Port, PACE, Linecard, CPU)
Destination (Port, PACE, Linecard, CPU)

Fig. 11

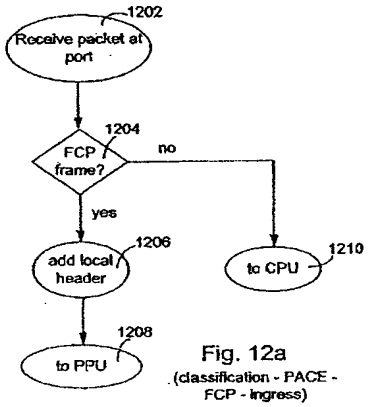


Fig. 12a
(classification - PACE - FCP - ingress)

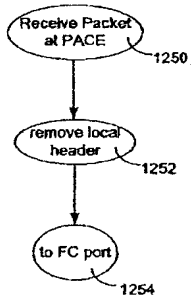


Fig. 12b
(classification - PACE - FCP - egress)

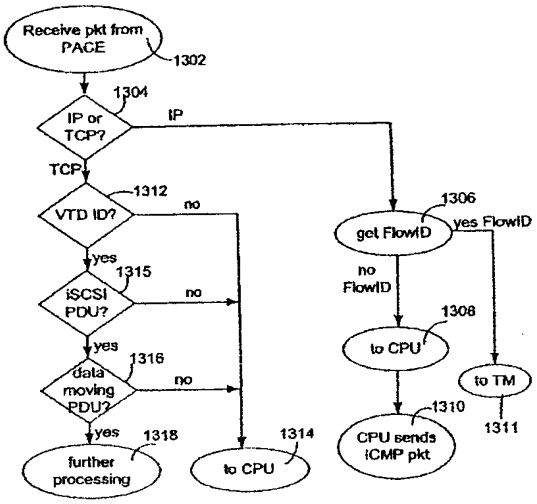


Fig. 13a
(Classification - PPU - ingress)

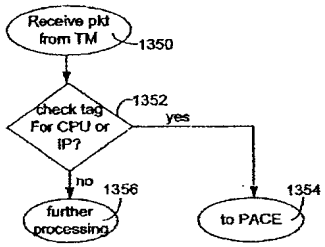


Fig. 13b
(Classification - PPU - egress)

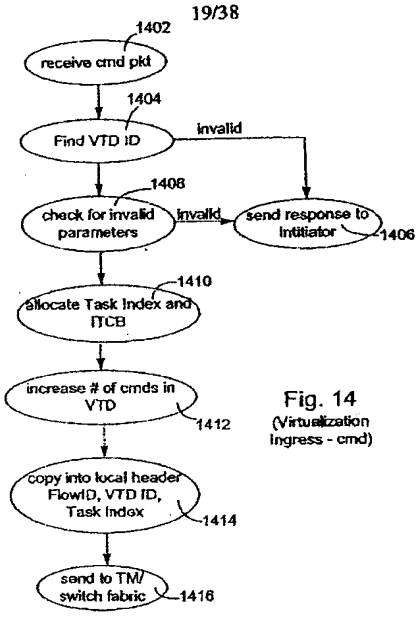


Fig. 14
(Virtualization Ingress - cmd)

20/38

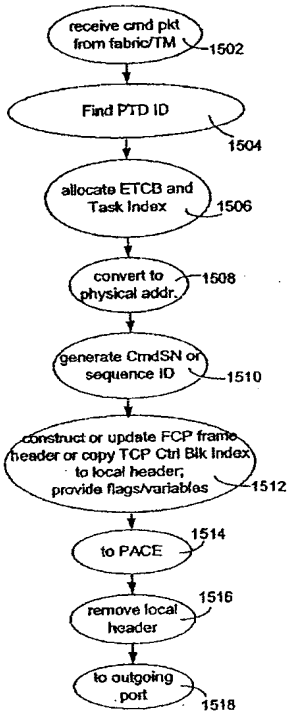


Fig. 15
(Virtualization - Egress - cmd)

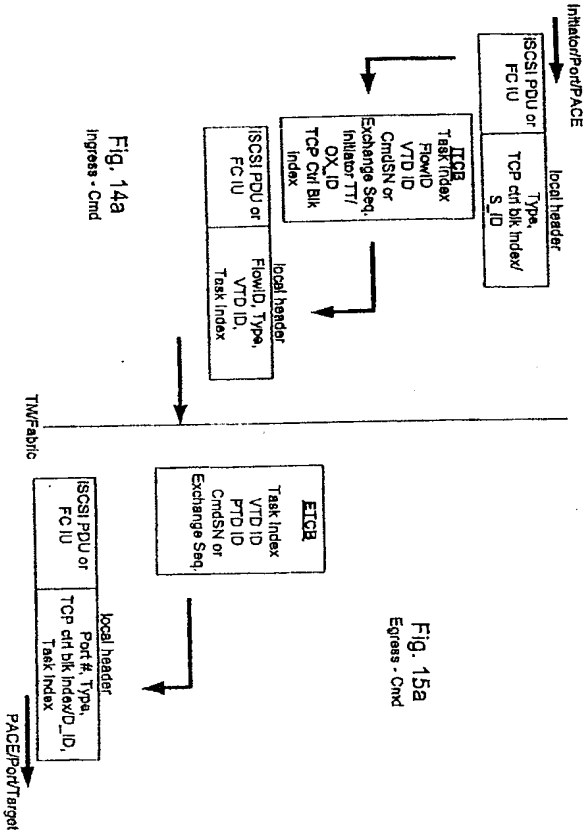


Fig. 14a
Ingress - Cmd

Fig. 15a
Egress - Cmd

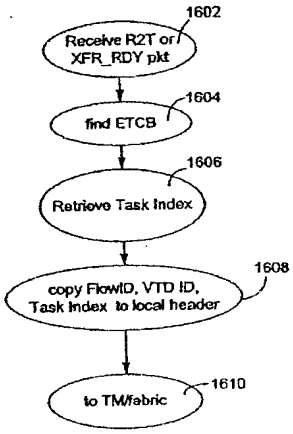


Fig. 16
(Virtualization - Ingress - R2T/XFR_RDY)

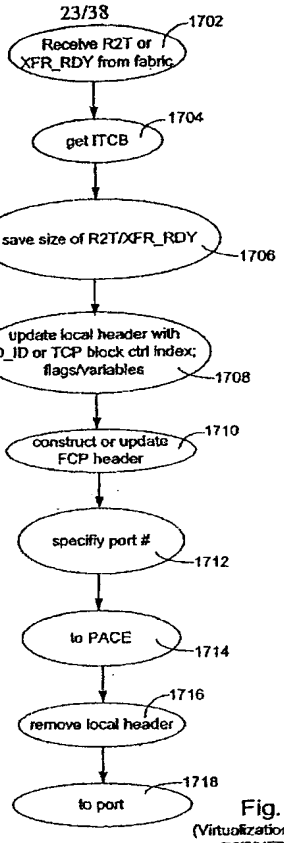


Fig. 17
(Virtualization - Egress - R2T/XFR_RDY)

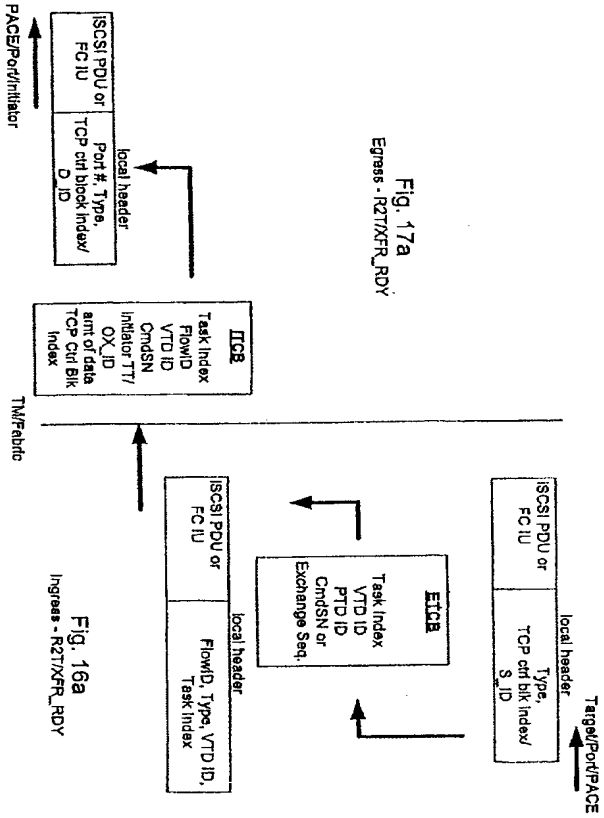


Fig. 17a
Egress - R2T/XFR_RDY

Fig. 16a
Ingress - R2T/XFR_RDY

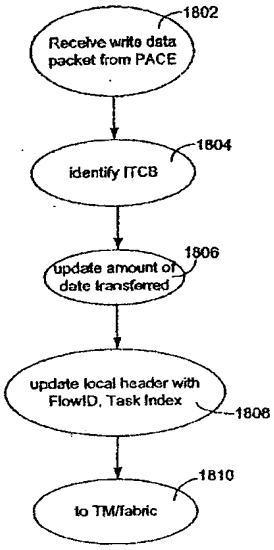


Fig. 18
(Virtualization - Ingress - write data packet)

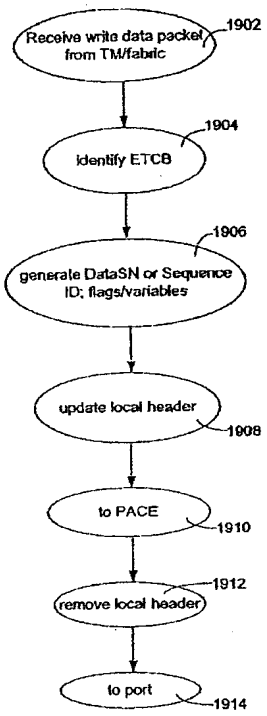


Fig. 19
(Virtualization - Egress - write data pkt)

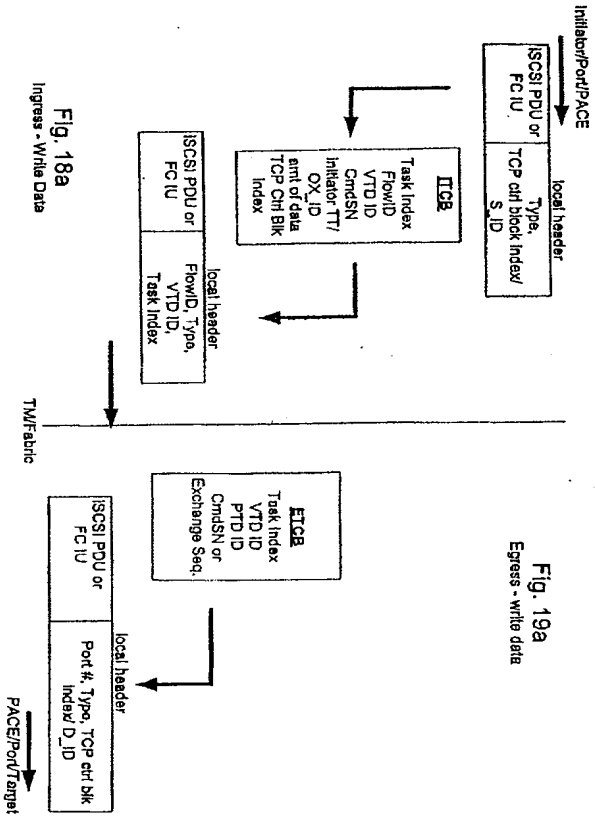


Fig. 18a
Ingress - Write Data

Fig. 19a
Egress - write data

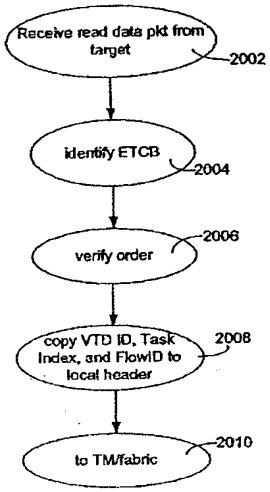


Fig. 20
(Virtualization - Ingress - Read Data pkt)

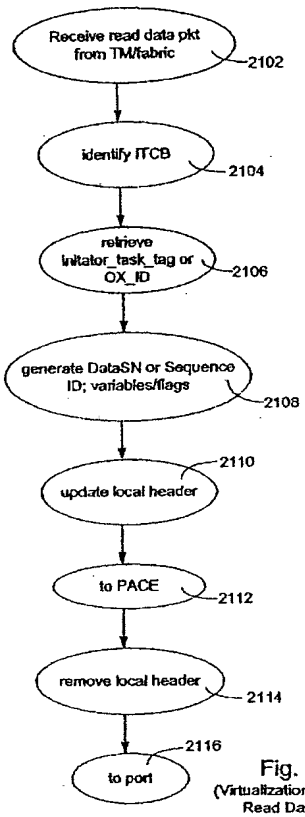


Fig. 21
(Virtualization - Egress-Read Data pkt)

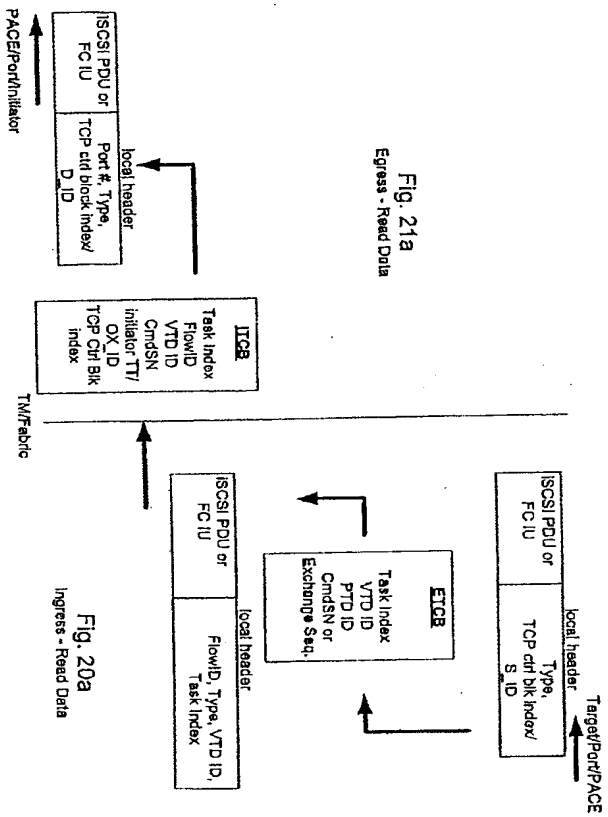


Fig. 21a
Egress - Read Data

Fig. 20a
Ingress - Read Data

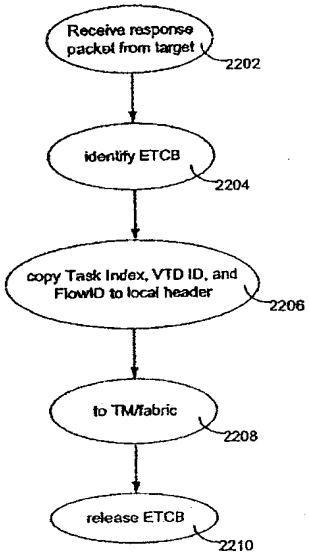


Fig. 22
(Virtualization - Ingress - response pkt)

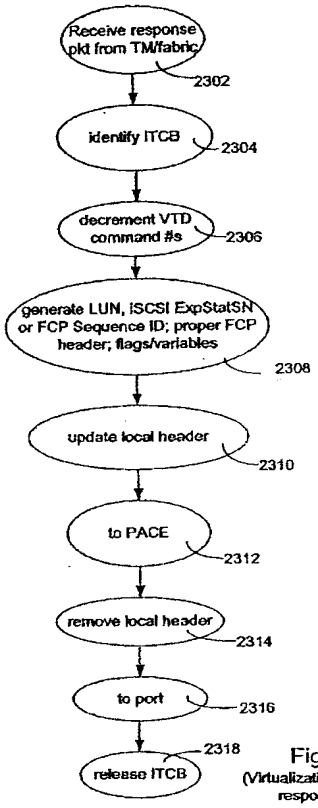


Fig. 23
(Virtualization - Egress - response pkt)

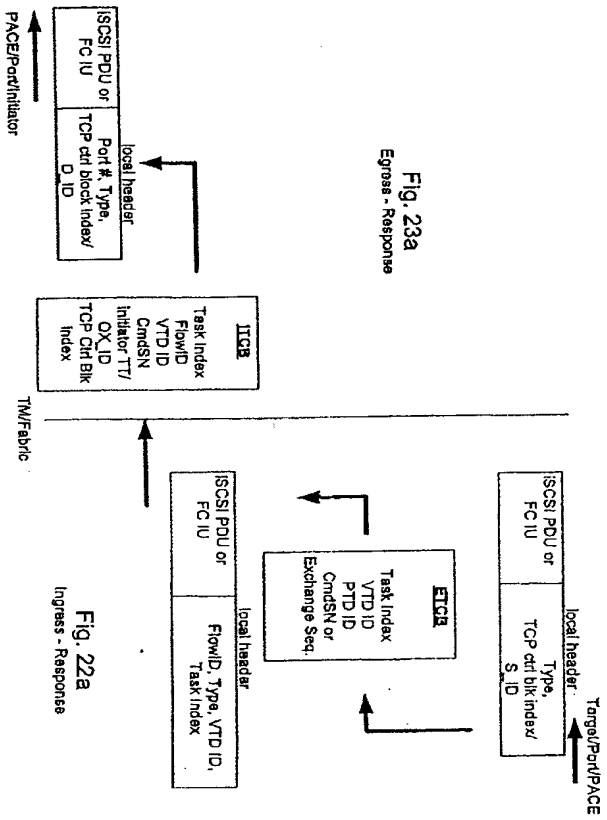


Fig. 23a
Egress - Response

Fig. 22a
Ingress - Response

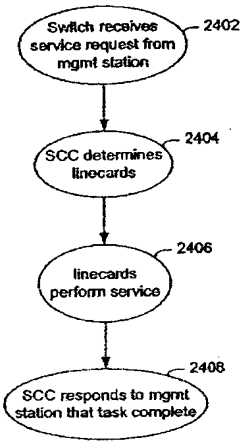

Switch receives service request from mgmt station
2402
SCC determines linecards
2404
linecards perform service
2406
SCC responds to mgmt station that task complete
2408

Fig.24

35/38

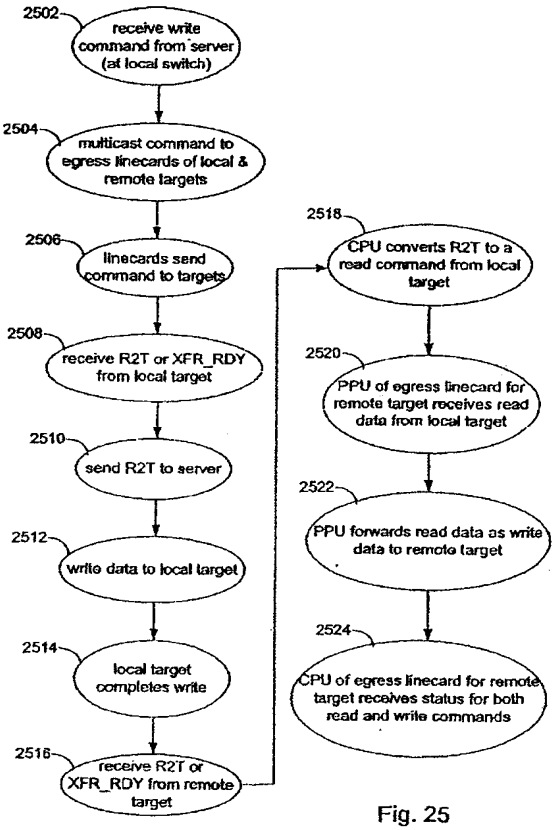

2502
receive write command from server (at local switch)
2504
multicast command to egress linecards of local & remote targets
2506
linecards send command to targets
2508
receive R2T or XFR_RDY from local target
2510
send R2T to server
2512
write data to local target
2514
local target completes write
2516
receive R2T or XFR_RDY from remote target
2518
CPU converts R2T to a read command from local target
2520
PPU of egress linecard for remote target receives read data from local target
2522
PPU forwards read data as write data to remote target
2524
CPU of egress linecard for remote target receives status for both read and write commands

Fig. 25

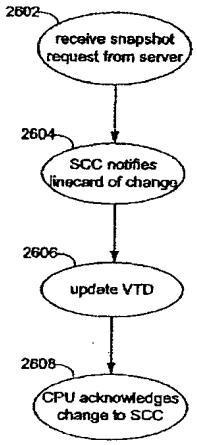
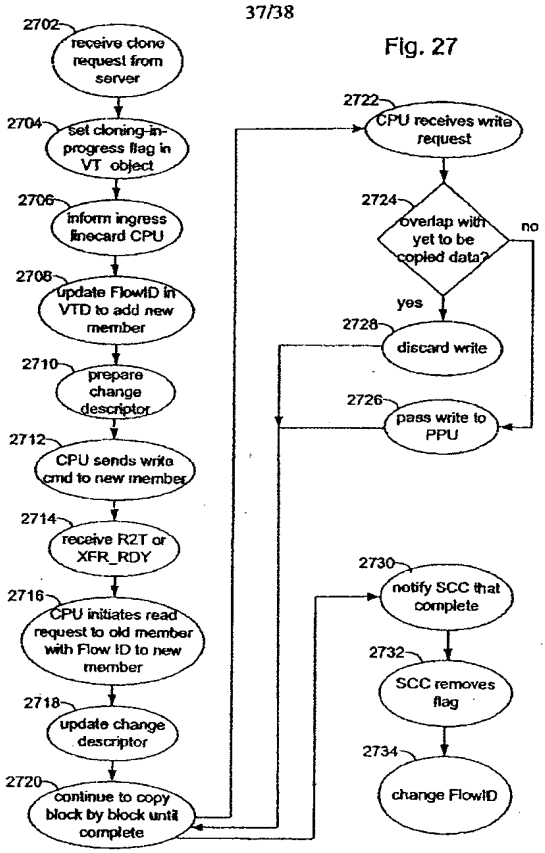

2602
receive snapshot request from server
2604
SCC notifies linecard of change
2606
update VTD
2608
CPU acknowledges change to SCC

Fig. 26

37/38

Fig. 27


2702
receive clone request from server
2704
set cloning-in-progress flag in VT object
2706
inform ingress linecard CPU
2708
update FlowID in VTD to add new member
2710
prepare change descriptor
2712
CPU sends write cmd to new member
2714
receive R2T or XFR_RDY
2716
CPU initiates read request to old member with Flow ID to new member
2718
update change descriptor
2720
continue to copy block by block until complete
2722
CPU receives write request
2724
overlap with yet to be copied data?
no
yes
2728
discard write
2726
pass write to PPU
2730
notify SCC that complete
2732
SCC removes flag
2734
change FlowID

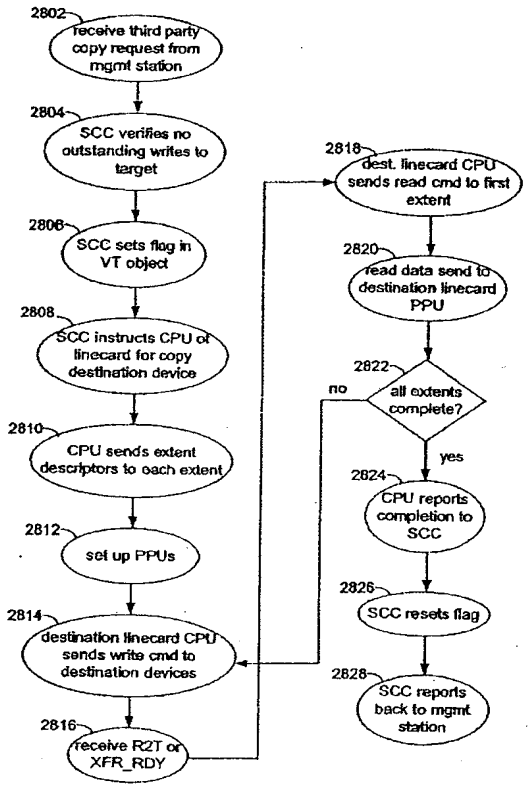

2802
receive third party copy request from mgmt station
2804
SCC verifies no outstanding writes to target
2806
SCC sets flag in VT object
2808
SCC instructs CPU of linecard for copy destination device
2810
CPU sends extent descriptors to each extent
2812
set up PPUs
2814
destination linecard CPU sends write cmd to destination devices
2816
receive R2T or XFR_RDY
2818
dest. linecard CPU sends read cmd to first extent
2820
read data send to destination linecard PPU
2822
all extents complete?
no
yes
2824
CPU reports completion to SCC
2826
SCC resets flag
2828
SCC reports back to mgmt station

Fig. 28

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/30912

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : G06F 15/16, G06F 9/00, G06F 15/177, H04L 12/56, H04J 3/16
US CL : 709/230, 220, 106, 200, 100, 101, 226-229, 250; 370/466, 465, 395, 402, 409, 355, 401
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 709/230, 220, 106, 200, 100, 101, 226-229, 250; 370/466, 465, 395, 402, 409, 355, 401

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
EAST: USPAT, US-PGPUB, EPO, JPO, IBM-TDB

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
fibre channel, iSCSI packet, format, protocol, translate, tcp

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y, P	U.S 6,600,730 B1 (Latif et al.) 04 June 2002. Abstract col. 1, lines 19-35, col. 2, line 15 - col. 4, line 49 col. 6, line 44 - col. 8, line 9	1-49
A	U.S 5,436,886 A (McGill) 25 July 1995 col. 3, line 57 - col. 5, line 12 col. 6, line 51 - col. 7, line 4	34, 37, 43, 48
Y, P	U.S 2002/0099797 A1 (Merrell et al.) 25 July 2002 Abstract pages 1-2: paragraphs 6, 10, pages 3-4, paragraphs 38	1-33, 35-36, 38-42, 44-47
Y, P	U.S 6,434,656 B1 (Downer et al.) 13 August 2002 col. 4, lines 15-49, col. 6, line 45 - col. 7, line 34	1-33
Y	U.S 6,493,750 B1 (Mathew et al.) 10 December 2002 Abstract, col. 1, lines 19-34, col. 2, lines 5-43, col. 4, lines 10-61, col. 5, line 63 - col. 8, line 36	1-49

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 January 2003 (23.01.2003)	28 FEB 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Mark Powell Telephone No. (703)305-9000

Form PCT/ISA/210 (second sheet) (July 1998)

フロントページの続き

(81)指定国 AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW

(72)発明者 ロレイアカー サントッシュ シー
アメリカ合衆国 カリフォルニア州 ９４０８６ サニーヴェイル ＃ジー ２０８ サウス フェア オークス アベニュー ６５５
(72)発明者 チェン ユ ピン
アメリカ合衆国 カリフォルニア州 ９５１２０ サン ホセ ギャレット コート １１７０
(72)発明者 リー ワン フイ
アメリカ合衆国 カリフォルニア州 ９５０３５ ミルピタス オークハースト ウェイ ３０３
(72)発明者 マラノン リネイト イー
アメリカ合衆国 カリフォルニア州 ９４５５０ リヴァーモア ノーマンディ サークル ２２４７
(72)発明者 タン エンニュー
アメリカ合衆国 カリフォルニア州 ９５１２８ サン ホセ フルートデイル アベニュー １９１９ アパートメント イー２１３

Fターム(参考) 5B014 EB05
5B065 BA01 CA18 CA30 CE21 ZA08
5B082 HA05 HA07
5K034 AA02 AA10 AA20 BB01 BB06 CC01 DD03 EE11 FF02 GG03
HH12 HH25 HH42 HH61 MM24